

A Faster Algorithm for Fully Dynamic Betweenness Centrality^{*}

Matteo Pontecorvi and Vijaya Ramachandran

University of Texas at Austin, USA
 {cavia,vlr}@cs.utexas.edu

Abstract. We present a new fully dynamic algorithm for maintaining betweenness centrality (BC) of vertices in a directed graph $G = (V, E)$ with positive edge weights. BC is a widely used parameter in the analysis of large complex networks. We achieve an amortized $O(\nu^{*2} \cdot \log^2 n)$ time per update, where $n = |V|$ and ν^* bounds the number of distinct edges that lie on shortest paths through any single vertex. This result improves on the amortized bound for fully dynamic BC in [21,22] by a logarithmic factor. Our algorithm uses new data structures and techniques that are extensions of the method in the fully dynamic algorithm in Thorup [28] for APSP in graphs with unique shortest paths. For graphs with $\nu^* = O(n)$, our algorithm matches the fully dynamic APSP bound in Thorup [28], which holds for graphs with $\nu^* = n - 1$, since it assumes unique shortest paths.

1 Introduction

Betweenness centrality (BC) is a widely-used measure in the analysis of large complex networks, and is defined as follows. Given a directed graph $G = (V, E)$ with $|V| = n$, $|E| = m$ and positive edge weights, let σ_{xy} denote the number of shortest paths (SPs) from x to y in G , and $\sigma_{xy}(v)$ the number of SPs from x to y in G that pass through v , for each pair $x, y \in V$. Then, $BC(v) = \sum_{s \neq v, t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}}$. As in [5], we assume positive edge weights to avoid the case when cycles of 0 weight are present in the graph.

The measure $BC(v)$ is often used as an index that determines the relative importance of v in G , and is computed for all $v \in V$. Some applications of BC include analyzing social interaction networks [13], identifying lethality in biological networks [20], and identifying key actors in terrorist networks [6,15]. In the static case, the widely used algorithm by Brandes [5] runs in $O(mn + n^2 \log n)$ on weighted graphs. Several approximation algorithms are available: [1,24] for static computation and, recently, [4,3] for dynamic computation. Heuristics for dynamic betweenness centrality with good experimental performance are given in [10,16,26], but none provably improve on Brandes. The only earlier exact dynamic BC algorithms that provably improve on Brandes on some classes of graphs are the recent separate incremental and decremental¹ algorithms in [18,19]. Recently, we give in [21] (see also [22]) the first fully dynamic algorithm for BC (the PR method) that is provably faster than Brandes for the class of dense graphs (where m is close to n^2) with succinct single-source SP dags. Table 1 contains a summary of these results.

In this paper, we present an improved algorithm for computing fully dynamic exact betweenness centrality: our algorithm FFD improves over PR by a logarithmic factor using data structures and technique that are considerably more complicated. Our method is a generalization of Thorup [28] (the Thorup method) which computes fully dynamic APSP for graphs with a unique SP for every vertex pair; however, a key step in BC algorithms is computing *all* SPs for each pair of vertices (*all pairs all shortest paths* — *APASP*). We develop a faster fully dynamic algorithm for APASP, which in turn leads to a faster fully dynamic BC algorithm than PR.

^{*} This work was supported in part by NSF grant CCF-1320675. Authors' affiliation: Dept. of Computer Science, University of Texas, Austin, TX 78712.

¹ Incremental/decremental refer to the insertion/deletion of a vertex or edge; the corresponding weight changes that apply are weight decreases/increases, respectively.

Our Results. Let ν^* be the maximum number of distinct edges that lie on shortest paths through any given vertex in G ; for convenience we assume $\nu^* = \Omega(n)$.

Theorem 1. *Let Σ be a sequence of $\Omega(n)$ fully dynamic vertex updates on a directed n -node graph $G = (V, E)$ with positive edge weights. Let ν^* bound the number of distinct edges that lie on shortest paths through any single vertex in any of the updated graphs or their vertex induced subgraphs. Then, all BC scores (and APASP) can be maintained in amortized time $O(\nu^{*2} \cdot \log^2 n)$ per update with algorithm FFD.*

Similar to the PR algorithm, our new algorithm FFD is provably faster than Brandes on dense graphs with succinct single-source SP dags. It also matches the Thorup bound for APSP when $\nu^* = O(n)$.

Our techniques rely on recomputing BC scores using certain data structures related to shortest paths extensions (see Section 2). These are generalizations of structures introduced by Demetrescu and Italiano [8] for fully dynamic APSP (the DI method) and Thorup, where only one SP is maintained for each pair of vertices. Our generalizations build on the tuple-system introduced in [19] (the NPRdec method) for decremental APASP (see Section 3), which is a method to succinctly represent all of the multiple SPs for every pair of vertices. Our algorithm also builds on the fully dynamic APASP (and BC) algorithm in PR, which runs in amortized $O(\nu^{*2} \cdot \log^3 n)$ cost per update (see Section 3). Finally, one of the main challenges we address in our current result is to generalize the ‘level graphs’ of Thorup to the case when different SPs for a given vertex pair can be distributed across multiple levels.

Paper	Year	Time	Weights	Update Type	DR/UN	Result
Brandes [5]	2001	$O(mn)$	NO	Static Alg.	Both	Exact
Brandes [5]	2001	$O(mn + n^2 \log n)$	YES	Static Alg.	Both	Exact
Geisberger et al. [9]	2007	Heuristic	YES	Static Alg.	Both	Approx.
Riondato et al. [24]	2014	depends on ϵ	YES	Static Alg.	Both	ϵ -Approx.
Semi Dynamic						
Green et al. [10]	2012	$O(mn)$	NO	Edge Inc.	Both	Exact
Kas et al. [12]	2013	Heuristic	YES	Edge Inc.	Both	Exact
NPR [18]	2014	$O(\nu^* \cdot n)$	YES	Vertex Inc.	Both	Exact
NPRdec [19]	2014	$O(\nu^{*2} \cdot \log n)$	YES	Vertex Dec.	Both	Exact
Bergamini et al. [4]	2015	depends on ϵ	YES	Batch (edges) Inc.	Both	ϵ -Approx.
Fully Dynamic						
Lee et al. [16]	2012	Heuristic	NO	Edge Update	UN	Exact
Singh et al. [26]	2013	Heuristic	NO	Vertex Update	UN	Exact
Kourtellis+ [14]	2014	$O(mn)$	NO	Edge Update	Both	Exact
Bergamini et al. [3]	2015	depends on ϵ	YES	Batch (edges)	UN	ϵ -Approx.
PR [21,22]	2015	$O(\nu^{*2} \cdot \log^3 n)$	YES	Vertex Update	Both	Exact
This paper (FFD)	2015	$O(\nu^{*2} \cdot \log^2 n)$	YES	Vertex Update	Both	Exact

Table 1. Related results (DR stands for Directed and UN for Undirected)

Discussion of the parameters m^* and ν^* . Let m^* be the number of distinct edges in G that lie on shortest paths; ν^* , defined above, is the maximum number of distinct edges on any single source SP dag. Clearly, $\nu^* \leq m^* \leq m$.

- m^* vs m : In many cases, $m^* \ll m$: as noted in [11], in a complete graph ($m = \Theta(n^2)$) where edge weights are chosen from a large class of probability distributions, $m^* = O(n \log n)$ with high probability.

- ν^* vs m^* : Clearly, $\nu^* = O(n)$ in any graph with only a constant number of SPs between every pair of vertices. These graphs are called k -geodetic [23] (when at most k SPs exists between two

nodes), and are well studied in graph theory [27,2,17]. In fact $\nu^* = O(n)$ even in some graphs that have an exponential number of SPs between some pairs of vertices.

In contrast, m^* can be $\Theta(n^2)$ even in some graphs with unique SPs, for example the complete unweighted graph K_n .

Another type of graph with $\nu^* \ll m^*$ is one with large clusters of nodes (e.g., as described by the *planted ℓ -partition model* [7,25]). Consider a graph H with k clusters of size n/k (for some constant $k \geq 1$) with $\delta < w(e) \leq 2\delta$, for some constant $\delta > 0$, for each edge e in a cluster; between the clusters is a sparse interconnect. Then $m^* = \Omega(n^2)$ but $\nu^* = O(n)$. (The connections between clusters can be arbitrary, thus BC scores are non-trivial to compute.) For all of the above classes of graphs, our BC algorithm will run in amortized $\tilde{O}(n^2)$ time per update (\tilde{O} hides polylog factors). More generally we have:

Theorem 2. *Let Σ be a sequence of $\Omega(n)$ updates on graphs with $O(n)$ distinct edges on shortest paths through any single vertex in any vertex-induced subgraph. Then, all BC scores (and APASP) can be maintained in amortized time $O(n^2 \cdot \log^2 n)$ per update.*

Our algorithm uses $\tilde{O}(m \cdot \nu^*)$ space, extending the $\tilde{O}(mn)$ result in DI for APSP. Brandes uses only linear space, but all known dynamic algorithms require at least $\Omega(n^2)$ space.

Overview of the Paper. In Section 2 we describe how to obtain a fully dynamic BC algorithm using a fully dynamic APASP algorithm. The remaining sections in the paper are devoted to developing our improved fully dynamic APASP algorithm. In Section 3 we review the **NPRdec**, **PR** and **Thorup** methods. In Section 4 we introduce the level tuple-system framework for APASP, with particular reference to the new data structures specifically developed for our result. In Section 5 we present our algorithm **FFD**, and we describe its main components in detail. Section 6 describes two important challenges that arise when generalizing **Thorup** to APASP setting; addressing these two challenges is crucial to the correctness and efficiency of our algorithm. Finally, in Section 7 we establish the amortized time bound of $O(\nu^{*2} \cdot \log^2 n)$ for **FFD** and its correctness.

2 Fully Dynamic Betweenness Centrality

The static Brandes algorithm [5] computes BC scores in a two phase process. The first phase computes the SP out-dag for every source through n applications of Dijkstra’s algorithm. The second phase uses an ‘accumulation’ technique that computes all BC scores using these SP dags in $O(n \cdot \nu^*)$ time.

In our fully dynamic algorithm, we will leave the second phase unchanged. For the first phase, we will use the approach in the incremental BC algorithm in [18], which maintains the SP dags using a very simple and efficient incremental algorithm. For decremental and fully dynamic updates the method is more involved, and dynamic APASP is at the heart of maintaining the SP dags. Neither the decremental nor our new fully dynamic APASP algorithms maintain the SP dags explicitly, instead they maintain data structures to update a collection of *tuples* (see Section 3). We now describe a very simple method to construct the SP dags from these data structures; this step is not addressed in the decremental **NPRdec** algorithm and it is only sketched in the fully dynamic **PR** algorithm.

For every vertex pair x, y , the following sets $R^*(x, y)$, $L^*(x, y)$ (introduced in DI) are maintained in **NPRdec**, and in both of our fully dynamic algorithms:

- $R^*(x, y)$ contains all nodes y' such that every shortest path $x \rightsquigarrow y$ in G can be extended with the edge (y, y') to generate another shortest path $x \rightsquigarrow y \rightarrow y'$.
- $L^*(x, y)$ contains all nodes x' such that every shortest path $x \rightsquigarrow y$ in G can be extended with the

edge (x', x) to generate another shortest path $x' \rightarrow x \rightsquigarrow y$.

These sets allow us to construct the SP dag for each source s using the following simple algorithm BUILD-DAG.

Algorithm 1 BUILD-DAG(G, s, \mathbf{w}, D) (\mathbf{w} is the weight function; D is the distance matrix)

```

1: for each  $t \in V$  do
2:   for each  $u \in R^*(s, t)$  do
3:     if  $D(s, t) + \mathbf{w}(t, u) = D(s, u)$  then add the edge  $(t, u)$  to dag( $s$ )

```

In our fully dynamic algorithm R^* and L^* will be supersets of the exact collections of nodes defined above, but the check in Step 3 will ensure that only the correct SP dag edges are included. The combined sizes of these R^* and L^* sets is $O(n \cdot \nu^* \cdot \log n)$ in our FFD algorithm, hence the amortized time bound for the overall fully dynamic BC algorithm is dominated by the time bound for fully dynamic APASP. In the rest of this paper, we will present our fully dynamic APASP algorithm FFD.

3 Background

In this section we review prior work upon which we build our results. For each, we highlight the inherited notation we use and the main ideas we extend.

3.1 The NPRdec Decremental APASP Algorithm [19]

NPRdec generalizes the decremental APSP algorithm in DI [8] to obtain a decremental algorithm for APASP and BC. For the decremental APSP algorithm DI develops a novel method to maintain (unique) shortest paths [8]. DI defines an *LSP* as a path where every proper subpath is a shortest path in the graph. By efficiently maintaining all LSPs after each update, DI presents an efficient decremental APSP, which is then extended to a fully dynamic APSP algorithm with additional tools. This provides a fully dynamic algorithm for APSP that runs in $O(n^2 \log n)$ amortized time per update. NPRdec extends this result to APASP by introducing the *tuple-system* to replace the need to maintain every SP and LSP in DI. We now briefly review this system, referring the reader to [19] for more details. Let $d(x, y)$ denote the shortest path length from x to y .

A *tuple*, $\tau = (xa, by)$, represents a set of paths in G , all with the same weight, and all of which use the same first edge (x, a) and the same last edge (b, y) . If the paths in τ are LSPs, then τ is an LST (locally shortest tuple), and the weight of every path in τ is $\mathbf{w}(x, a) + d(a, b) + \mathbf{w}(b, y)$. In addition, if $d(x, y) = \mathbf{w}(x, a) + d(a, b) + \mathbf{w}(b, y)$, then τ is a *shortest path tuple* (ST).

A *triple* $\gamma = (\tau, wt, count)$, represents the tuple $\tau = (xa, by)$ that contains *count* number of paths from x to y , each with weight *wt*. NPRdec uses triples to succinctly store all LSPs and SPs for each vertex pair in G .

Left Tuple and Right Tuple. A left tuple (or ℓ -tuple), $\tau_\ell = (xa, y)$, represents the set of LSPs from x to y , all of which use the same first edge (x, a) . A right tuple (r -tuple) $\tau_r = (x, by)$ is defined analogously. For a shortest path r -tuple $\tau_r = (x, by)$, $L(\tau_r)$ is the set of vertices which can be used as pre-extensions to create LSTs in G , and for a shortest path ℓ -tuple $\tau_\ell = (xa, y)$, $R(\tau_\ell)$ is the set of vertices which can be used as post-extensions to create LSTs in G .

NPRdec maintains several other sets such as P^* and P for each vertex pair. Since our algorithm also maintains generalizations of these sets, we will discuss them when we present the data structures used by our algorithm in Section 4.

Similar to DI, the **NPRdec** algorithm first deletes all the paths containing the updated node using a procedure **CLEANUP**, and then updates the tuple-system to maintain all shortest paths in the graph using procedure **FIXUP**. The main difference, in terms of data structures, is the use of tuples to collect paths that share the same first edge, last edge and weight.

3.2 The PR Fully Dynamic APASP Algorithm [21]

In our recent fully dynamic PR algorithm for APASP [21], we build on the tuple-system introduced in **NPRdec**. PR also incorporates several elements in the DI method of extending their decremental APSP algorithm to fully dynamic (though some key elements in PR are significantly different from DI).

One difference between **NPRdec** and PR is the introduction of HTs and LHTs; these are extension of historical and locally historical paths in DI to tuples and triples defined as follows (THTs and TLHTs are not specifically used in this paper, however we include their definitions for completeness):

HT, THT, LHT, and TLHT. Let τ be a tuple in the tuple-system at time t . Let $t' \leq t$ denote the most recent step at which a vertex on a path in τ was updated. Then τ is a *historical tuple (HT)* at time t if τ was an ST-tuple at least once in the interval $[t', t]$; τ is a *true HT (THT)* at time t if it is not an ST in the current graph. A tuple τ is a *locally historical tuple (LHT)* at time t if either it only contains a single vertex or every proper sub-path in it is an HT at time t ; a tuple τ is a *true LHT (TLHT)* at time t if it is not an LST in the current graph.

Similar to DI for unique SPs, PR forms LHTs and TLHTs in its **FULLY-FIXUP** procedure (which adapts the **FIXUP** procedure in **NPRdec** to the fully dynamic case) by combining compatible pairs of HTs. An important method introduced in DI for efficiency in fully dynamic APSP is the notion of a ‘dummy update’ sequence. Extending this method to an efficient algorithm for APASP does not appear to be feasible, so in PR, a new dummy update sequence (that uses elements in a different method by Thorup [28]) is introduced. PR then defines the *Prior Deletion Graphs* (PDGs) (that are related to the *level graphs* maintained in Thorup – see below) to study the complexity of the PR algorithm. In our algorithm **FFD** we will use graphs similar to the PDGs in PR; these are described in Section 4.

3.3 The Thorup Fully Dynamic APSP Algorithm [28]

In [28], Thorup improves by a logarithmic factor over DI (for unique shortest paths) by using a *level system* of decremental only graphs. The shortest paths and locally shortest paths are generated level by level leading to a different complexity analysis from DI. When a node is removed from the current graph, it is also removed from every older level graph that contains it. The implementation of the Thorup APSP algorithm is not fully specified in [28]. For our **FFD** algorithm, we present generalizations of the data structures sketched in Thorup (see Section 4 for a summary of these data structures).

4 Data Structures for Algorithm **FFD**

Our algorithm **FFD** requires several data structures. Some of these are already present in **NPRdec**, PR and Thorup, while others are newly defined or generalized from earlier ones. We will use components from PR such as the abstract representation of the level system using PDGs (see Section 4) and

the flag bit β for a triple, the Marked-Tuples scheme introduced in **NPRdec** (see [19,21] for more details), and the maintenance of level graphs from **Thorup**.

In this section, we describe all data structures used by our algorithm. In Table 2 we summarize the structures we use, including those inherited from [19,21,28]. The new components we introduce in this paper to achieve efficiency for fully dynamic APASP, are described in section 4.2 and listed in Table 2, Part D.

4.1 A Level System for Centered Tuples

Algorithm FFD uses the PDGs defined in **PR** as real data structures similar to **Thorup** for APSP. This is done in order to generate a smaller superset of LSTs than **PR**, and this is the key to achieving the improved efficiency. Here we describe the level system and the data structures we use in FFD, with special attention to the new elements we introduce.

As in [8,21,28] we build up the tuple-system for the initial n -node graph $G = (V, E)$ with n insert updates (starting with the empty graph), and we then perform n updates according to the update sequence Σ . After $2n$ updates, we reset all data structures and start afresh.

Our level system is a generalization of **Thorup** to fully dynamic APASP. For an update at step t , let k be the position of the least significant bit with value 1 in the binary representation of t . Then the t -th update activates the level k , and deactivates all levels $j < k$ by folding these levels into level k . These levels are considered implicitly in **PR**, and using the same notation, we will say that $time(k) = t$, and $level(t) = k$; moreover G_t indicates the graph after the t -th update. Note that the largest level created before we start afresh is $r = \log 2n$.

Centering vertices and tuples/triples. As in **Thorup**, each *vertex* v is *centered* in level $k = level(t)$, where t is the most recent step in which v was updated. A path p in a tuple is centered in level $k' = level(t')$, where t' is the most recent step in which p entered the tuple system (within some tuple) or was modified by a vertex update. Hence, in contrast to **Thorup**, a triple can represent paths centered in different levels. Thus, for a triple $\gamma = ((xa, by), wt, count)$ we maintain an array C_γ where

$$C_\gamma[i] = \text{number of paths represented by } \gamma \text{ that are all centered in level } i$$

It follows that $\sum_i C_\gamma[i] = count$. The *level center of the triple* γ is the smallest (i.e., most recent level) i such that $C_\gamma[i] \neq 0$.

Level graphs (PDGs). The **PR** algorithm defines PDGs as follow: Let t be the current update step, let $t' < t$, and let W be the set of vertices that are updated in the interval of steps $[t' + 1, t]$. The *prior deletion graph (PDG)* $\Gamma_{t'}$ is the induced subgraph of $G_{t'}$ on the vertex set $V(G_{t'}) - W$.

In **PR**, the PDGs are used only in the analysis, and are not maintained by the algorithm. Here, in FFD, we will maintain a set of local data structures for each PDG that is relevant to the current graph; also, in a small change of notation, we will denote a level graph for time $t' \leq t$ as $\Gamma_{k'}$, where $k' = level(t')$ rather than the **PR** notation of $\Gamma_{t'}$. These graphs are similar to the level graphs in **Thorup**. As in **Thorup**, only certain information for Γ_k is explicitly maintained in its local data structures: the STs centered in level k plus all the extensions that can generate STs in Γ_k . The data structures used by our algorithm to maintain triples are Global and Local, which we now describe.

Global Structures. The global data structures are P^* , P , L and R (see Table 2, Part A).

Notation	Data Structure	Appears
Part A :: Global Data Structures (for each pair of nodes (x, y))		
$P(x, y)$	all (centered) LHT s from x to y with weight as key	([8] for paths, [19] for LSTs)
$P^*(x, y)$	all (centered) HT s from x to y with weight as key	
$L(x, by)$	$\{x' : (x'x, by) \text{ denotes a (centered) LHT}\}$	[21], ([19] for LSTs)
$R(xa, y)$	$\{y' : (xa, yy') \text{ denotes a (centered) LHT}\}$	[21], ([19] for LSTs)
Marked-Tuples	global dictionary for Marking scheme	[21,19]
Part B :: Local Data Structures (for each active level i , for each pair of nodes (x, y))		
$P_i^*(x, y)$	STs from x to y centered in level i	(sketched in [28] for paths)
$L_i^*(x, y)$	$\{x' : (x'x, y) \text{ denotes an } \ell\text{-tuple for SPs centered in level } i\}$	
$R_i^*(x, y)$	$\{y' : (x, yy') \text{ denotes an } r\text{-tuple for SPs centered in level } i\}$	
$LC_i^*(x, y)$	the subset $\{x' \in L_i^*(x, y) : x' \text{ is centered in level } i\}$	
$RC_i^*(x, y)$	the subset $\{y' \in R_i^*(x, y) : y' \text{ is centered in level } i\}$	
$dict_i$	dictionary of pointers from local STs to global P and P^*	new
Part C :: Inherited Data Structures		
$\beta(\gamma)$	flag bit for the (centered) triple γ	[21]
$level(t)$	level activated during t -th update	[21]
$time(k)$	most recent update in which level k is activated	[21]
\mathcal{N}	nodes (centered in levels) deactivated in the current step	[21]
Γ_k	level graph (PDG) created during $time(k)$ -th update	([21] as PDG, [28] for paths)
Part D :: New Data Structures		
C_γ	distribution of all paths in triple γ among active levels	new
$DL(x, y)$	linked-list containing the history of distances from x to y	new
$LN(x, y, wt)$	the set $\{b : \exists (xa, by) \text{ of weight } wt \text{ in } P(x, y)\}$	new
$RN(x, y, wt)$	the set $\{a : \exists (xa, by) \text{ of weight } wt \text{ in } P(x, y)\}$	new

Table 2. Notation summary

- The structures $P^*(x, y)$ and $P(x, y)$ will contain HTs (including all STs) and LHTs (including all LSTs), respectively, from x to y . They are priority queues with the weights of the triple and a flag bit β as key. For a triple γ in P , the flag bit $\beta(\gamma) = 0$ if the triple γ is in P but not in P^* , and $\beta(\gamma) = 1$ if the triple γ is in P and P^* .
- The structure $L(x, by)$ ($R(xa, y)$) is the set of all left (right) extension vertices that generate a centered LHT in the tuple-system.

Local Structures. The local data structures we introduce in this paper are L_i^* , R_i^* , LC_i^* and RC_i^* (see Table 2, Part B). These are generalization of the data structures sketched in **Thorup** for unique SPs in the graph. For every pair of nodes (x, y) :

- The structure $P_i^*(x, y)$ contains the set of STs from x to y centered in Γ_i . It is implemented as a set.
- The structure $L_i^*(x, y)$ ($R_i^*(x, y)$) contains all left (right) extensions that generate a shortest ℓ -tuple (r -tuple) centered in level i . It is implemented as a balanced search tree.
- The structure $LC_i^*(x, y)$ ($RC_i^*(x, y)$) contains left (right) extensions centered in level i that generate a shortest ℓ -tuple (r -tuple) centered in level i . It is implemented as a balanced search tree.
- A dictionary $dict_i$, contains STs in P_i^* using the key $[x, y, a, b]$ and two pointers stored along with each ST. The two pointers refer to the location in $P(x, y)$ and $P^*(x, y)$ of the triple of the form (x, a, b, y) contained in $P_i^*(x, y)$.

In order to recompute BC scores (see Section 2) we will consider $R^*(x, y) = \bigcup_i R_i^*(x, y)$ and similarly $L^*(x, y) = \bigcup_i L_i^*(x, y)$.

4.2 New Structures

We introduce two completely new data structures which are essential to achieve efficiency for our FFD algorithm. Both are needed to address the Partial Extension Problem (PEP) which does not appear in **Thorup** (see section 6).

Distance History Matrix. The *distance history matrix* is a matrix DL where each entry is a pointer to a linked list: for each $x, y \in V$, the linked list $DL(x, y)$ contains the sequence of different pairs (wt, k) , where each one represents an SP weight wt from x to y , along with the most recent level k in which the weight wt was the shortest distance from x to y in the graph Γ_k . The pair with weight wt in $DL(x, y)$ is double-linked to every triple from x to y with weight wt in the system. Precisely, when a new triple γ from x to y of weight wt is inserted in the algorithm, a link is formed between γ and the pair (wt, k) in $DL(x, y)$. With this structure, the FFD algorithm can quickly check if there are still triples of a specific weight in the tuple-system, especially for example when we need to remove a given weight from DL (e.g. Step 21 - Alg. 5). Note that the size of each linked list is $O(\log n)$.

Historical Extension (HE) Sets RN and LN . Another important type of structure we introduce are the sets RN and LN . These structures are crucial to select efficiently the set of restored historical tuples that need to be extended (see Section 5.2). $RN(x, y, wt)$ ($LN(x, y, wt)$ works symmetrically) contains all nodes b such that there exists at least one tuple of the form $(x \times, by)$ and weight wt in $P(x, y)$. Similarly to DL , every time a new triple γ of this form is inserted in the tuple-system, a double link is created between γ and the occurrence of b in $RN(x, y, wt)$ in order to quickly access the triple when needed (e.g. Steps 14 to 16 - Alg. 9).

The total space used by DL , RN and LN is $O(n^2 \log n)$. This is dominated by the overall space used by the algorithm to maintain all the triples in the tuple-system across all levels (see Lemma 4, Section 7).

5 The FFD Algorithm

Algorithm FFD is similar to the fully dynamic algorithm in **PR** and its overall description is given in Algorithm 2. The main difference is the introduction of the notion of levels as described in Section 4, and their activation/deactivation as in **Thorup**. At the beginning of the t -th update (with $k = level(t)$), we first activate the new level k and we perform FF-UPDATE (Alg. 3) on the updated node v . As in **PR** and shown in Table 2 - Part C, the set \mathcal{N} consists of all vertices centered at these lower deactivated levels. All vertices in \mathcal{N} are re-centered at level k during the t -th update (Alg. 2, Step 5), and ‘dummy’ update operations are performed on each of these vertices. Note that \mathcal{N} contains the $2^k - 1$ most recently updated vertices in reverse order of update time (from the most recent to the oldest). Procedure FF-UPDATE is invoked with the parameter k representing the newly activated level. Finally, all levels $j < k$ are deactivated (Alg. 2, Step 6).

FF-UPDATE. As in [8,19,21], the update of a node occurs in a sequence of two steps: a *cleanup phase* and a *fixup phase*. Here, we call this update FF-UPDATE and it is a sequence of 2 calls: FF-CLEANUP and FF-FIXUP. Briefly, FF-CLEANUP removes all LHPs in the tuple-system containing the updated node (see Section 5.1), while FF-FIXUP identifies and adds the STs and LSTs in the updated graph

Algorithm 2 FFD(G, v, \mathbf{w}', k)

```
1: activate the new level  $k$ 
2: FF-UPDATE( $v, \mathbf{w}', k$ )
3: generate the set  $\mathcal{N}$ 
4: for each  $u \in \mathcal{N}$  in decreasing order of update time do
5:   FF-UPDATE( $u, \mathbf{w}', k$ ) {dummy updates}
6: deactivate all levels lower than  $k$ 
```

that are not yet in the tuple-system (see Section 5.2). Both FF-CLEANUP and FF-FIXUP are more involved algorithms than their counterparts in PR and the resulting algorithm will save a $O(\log n)$ factor over the amortized cost in PR.

Algorithm 3 FF-UPDATE(v, \mathbf{w}', k)

```
1: FF-CLEANUP( $v$ )
2: FF-FIXUP( $v, \mathbf{w}', k$ )
```

Dummy Updates. Calling FF-UPDATE only on the updated nodes gives us a correct fully dynamic APASP algorithm. However, the number of LHTs generated could be very large making this strategy not efficient in general. Thus, as in PR, the FFD algorithm performs a sequence of dummy updates as follows. Consider an update on v at time t . Let $k = \text{level}(t)$. As shown in algorithm 2, \mathcal{N} is the set of $2^k - 1$ most recently updated nodes $v_{t-1}, v_{t-2}, \dots, v_j$, where $j = 2^k - 1$. These are the nodes centered at levels smaller (more recent) than k before the t -th update is applied. For each vertex u in \mathcal{N} , starting from the most recent to the oldest, FFD calls FF-UPDATE on the node u . This procedure removes all LHTs containing u in every active level using FF-CLEANUP, and immediately reinserts them in the newly activated level k by performing FF-FIXUP and using the local data structures. These are the dummy updates. Dummy updates have the effect of removing from the tuple-system any path π containing a vertex in \mathcal{N} that is no longer an LSP in the current graph, because π is removed by the corresponding FF-CLEANUP step and will not be restored during FF-FIXUP.

We will establish in Section 7 that FFD correctly updates the data structures with the amortized bound given in Theorem 1.

5.1 Description of FF-CLEANUP

FF-CLEANUP removes all the LHPs going through the updated vertex v from all the global structures P, P^*, L and R , and from all local structures in any active level graph Γ_j that contains these triples. This involves decrementing the count of some triples or removing them completely (when all the paths in the triple go through v). The algorithm also updates local dictionaries and the DL, RN and LN structures. Algorithm FF-CLEANUP is a natural extension of the NPRdec cleanup. An extension of the NPRdec cleanup is used also in PR but in a different way.

Algorithm 4 FF-CLEANUP(v)

```
1:  $H_c \leftarrow \emptyset$ ; Marked-Tuples  $\leftarrow \emptyset$ 
2:  $\gamma \leftarrow [(v, v), 0, 1]$ ;  $C_\gamma[\text{center}(v)] = 1$ ; add  $[\gamma, C_\gamma]$  to  $H_c$ 
3: while  $H_c \neq \emptyset$  do
4:   extract in  $S$  all the triples with the same min-key  $[wt, x, y]$  from  $H_c$ 
5:   FF-CLEANUP- $\ell$ -extend( $S, [wt, x, y]$ ) (see Algorithm 5)
6:   FF-CLEANUP- $r$ -extend( $S, [wt, x, y]$ )
```

FF-CLEANUP starts as in the NPRdec algorithm. We add the updated node v to H_c (Step 2 – Alg. 4) and we start extracting all the triples with same min-key (Step 4 – Alg. 4). The main differences from NPRdec start after we call Algorithm 5. As in [19], we start by forming a new triple γ' to be deleted (Steps 5 – Alg. 4). A new feature in Algorithm 5 is to accumulate the paths that we need to remove level by level using the array C' . This is inspired by Thorup where unique SPs are maintained in each level. However, our algorithm maintains multiples paths spread across different levels using the C_γ arrays associated to LSTs, and the technique used to update the C_γ arrays is significantly different and more involved than the one described in Thorup. Step 6 - Alg. 5 calls FF-CLEANUP-CENTERS (Alg. 6) that will perform this task.

Algorithm 5 FF-CLEANUP- ℓ -extend($S, [wt, x, y]$)

```

1: for every  $b$  such that  $(x \times, by) \in S$  do
2:   let  $S_b \subseteq S$  be the set of all triples of the form  $(x \times, by)$ 
3:   let  $fcnt'$  be the sum of all the counts of all triples in  $S_b$ 
4:   for every  $x'$  in  $L(x, by)$  s.t.  $(x'x, by) \notin \text{Marked-Tuples}$  do
5:      $wt' \leftarrow wt + \mathbf{w}(x', x)$ ;  $\gamma' \leftarrow ((x'x, by), wt', fcnt')$ 
6:      $C_{\gamma'} \leftarrow \text{FF-CLEANUP-CENTERS}(\gamma', S_b)$ 
7:     add  $[\gamma', C_{\gamma'}]$  to  $H_c$ 
8:     remove  $\gamma'$  in  $P(x', y)$  // decrements count by  $fcnt'$ 
9:     set new center for  $\gamma'' = ((x'x, by), wt')$  in  $P(x', y)$  as  $\text{argmin}_i(C_{\gamma''}[i] \neq 0)$ 
10:    if a triple for  $(x'x, by)$  exists in  $P(x', y)$  then
11:      insert  $(x'x, by)$  in Marked-Tuples
12:    else
13:      delete  $x'$  from  $L(x, by)$  and delete  $y$  from  $R(x'x, b)$ 
14:    if no triple for  $((x'x, by), wt')$  exists in  $P(x', y)$  then
15:      remove  $b$  from  $RN(x', y, wt')$ 
16:    if no triple for  $((x'x, -y), wt')$  exists in  $P(x', y)$  then
17:      remove  $x$  from  $LN(x', y, wt')$ 
18:    if a triple for  $((x'x, by), wt')$  exists in  $P^*(x', y)$  then
19:      remove  $\gamma'$  in  $P^*(x', y)$  // decrements count by  $fcnt'$ 
20:      if  $\gamma' \notin P^*(x', y)$  then
21:        remove the element with weight  $wt'$  from  $DL(x', y)$  if not linked to other tuples in  $P^*(x', y)$ 
22:      for each  $i$  do
23:        decrement  $C_{\gamma'}[i]$  paths from  $\gamma' \in P_i^*(x', y)$ 
24:        if  $\gamma'$  is removed from  $P_i^*(x', y)$  then
25:          if  $x'$  is centered in level  $i$  then
26:            if  $\forall j \geq i, P_j^*(x, y) = \emptyset$  then
27:              remove  $x'$  from  $L_i^*(x, y)$  and remove  $x'$  from  $LC_i^*(x, y)$ 
28:            else if  $P_i^*(x, y) = \emptyset$  then
29:              remove  $x'$  from  $L_i^*(x, y)$ 
30:          if  $y$  is centered in level  $i$  then
31:            if  $\forall j \geq i, P_j^*(x', b) = \emptyset$  then
32:              remove  $y$  from  $R_i^*(x', b)$  and remove  $y$  from  $RC_i^*(x', b)$ 
33:            else if  $P_i^*(x', b) = \emptyset$  then
34:              remove  $y$  from  $R_i^*(x', b)$ 

```

FF-CLEANUP-CENTERS takes as input the generated triple γ' of the form $(x'x, by)$ which contains all the paths going through the updated node v to be removed, and the set of triples S_b of the form $(x \times, by)$ that are extended to x' to generate γ' . This procedure has two tasks: (1) generating the $C_{\gamma'}$ vector for the triple γ' that will be reinserted in H_c for further extensions, and (2) updating the $C_{\gamma''}$ vector for the tuple γ'' in $P(x', y)$ (note that γ'' is the corresponding triple in P of γ' , before we subtract all the paths represented by γ' level by level).

(1) - This task, which is more complex than the second task (which is a single step in the algorithm, see point (2) below), is accomplished in steps 4 to 8, Alg. 6 and uses the following technique. In step 4 – Alg. 6, we store into the $\log n$ -size array C' the distribution over the active levels for the set of triples in S_b that generates γ' using the left extension to x' . In order to generate the correct

Algorithm 6 FF-CLEANUP-CENTERS(γ', S_b)

```
1: let  $\gamma' = ((x'x, by), wt', fcount')$  (the triple of the form  $((x'x, by), wt')$  that contains all the paths through  $v$  to be removed)
2: let  $\gamma'' = ((x'x, by), wt'', fcount'')$  (the triple of the form  $((x'x, by), wt')$  in  $P(x', y)$ . Note that  $\gamma'$  represents a subset of  $\gamma''$ )
3:  $j \leftarrow \text{argmax}_j (C_{\gamma''}[j] \neq 0)$  // This is the oldest level in which a path in  $\gamma''$  appeared for the first time
4:  $C' \leftarrow \sum_{\gamma \in S_b} C_{\gamma}[r-1, \dots, 0]$  // This is the sum (level by level) of the triples of the form  $(xa_i, by)$  that go through  $v$  and
   are extending to  $\gamma'$  during this stage
5: create a new center vector  $C_{\gamma'}$  for the triple  $\gamma'$  as follows
6:   for all the levels  $m > j$  we set  $C_{\gamma'}[m] = 0$ 
7:   for the level  $j$  we set  $C_{\gamma'}[j] = \sum_{k=j}^{r-1} C'[k]$ 
8:   for all the levels  $i < j$  we set  $C_{\gamma'}[i] = C'[i]$ 
9:  $C_{\gamma''}[r-1, \dots, 0] \leftarrow C_{\gamma''}[r-1, \dots, 0] - C_{\gamma'}[r-1, \dots, 0]$  // We update the  $C$  vector for  $\gamma'' \in P(x', y)$ 
10: return  $C_{\gamma'}$  // We return the correct vector for the generated  $\gamma'$  triples
```

vector $C_{\gamma'}$ (to associate with the triple γ'), we need to reshape the distribution in C' according to the corresponding distribution of the triple $\gamma'' \in P$. The reshaping procedure works as follows: we first identify the oldest level j in which the triple γ'' appeared in P for the first time (Step 3 – Alg. 6). Recall that we want to remove γ' paths containing v from γ'' , and γ'' does not exist in any level older than j . Vector C' is the sum of C_{γ} for all $\gamma \in S_b$ (Step 4 – Alg. 6). Those triples are of the form (xa_i, by) and they could exist in levels older, equal or more recent than j . But the triples in S_b that were present in a level older than j , were extended to γ'' in P for the first time in level j . For this reason, step 7 - Alg. 6 aggregates all the counts in C' in levels older or equal j in $C_{\gamma'}[j]$. Moreover, for each level $i < j$, if a triple $\gamma \in S_b$ is present in the level graph Γ_i with *count* paths centered in level i , then Γ_i also contains its extension to x' that is a subtriple of γ'' located in level i with at least *count* paths. Thus for each level $i < j$ step 8 - Alg. 6, copies the number of paths level-wise. This procedure allows us to precisely remove the LHPs only from the level graphs where they exist. After C' is reshaped into $C_{\gamma'}$ (steps 5 to 8 - Alg. 6), the algorithm returns this correct array for γ' to Alg. 5.

(2) - This task is performed by the simple step 9, Alg. 6, which is a subtraction level by level of LHPs.

After adding the new triple γ' to H_c (Step 7 - Alg. 5), the algorithm continues as the **NPRdec** (Steps 7 to 13 – Alg. 5) with some differences: we need to update centers, local data structures, DL , RN and LN . We update the center of γ' using $C_{\gamma'}$ (Step 9 - Alg. 5). If γ' is a shortest triple, we decrement the count of $\gamma' \in P^*(x', y)$ (Step 19 - Alg. 5). If γ' is completely removed from $P^*(x', y)$ and $DL(x', y, wt')$ is not linked to any other tuple, we remove the entry with weight wt' from $DL(x', y)$ (Step 21 - Alg. 5). Moreover, we subtract the correct number of paths from each level using the (previously built) array $C_{\gamma'}$ (Step 23 - Alg. 5). Finally for each active level i , if γ' is removed from $P_i^*(x', y)$, we take care of the sets L_i^* and R_i^* (Steps 24 to 34 - Alg. 5). In the process, we also update LC_i^* and RC_i^* in case the endpoints of γ' are centered in level i . If γ' is completely removed from $P(x', y)$, using the double links to the node b in $RN(x', y, wt')$, we check if there are other triples that use b in $P(x', y)$ (Step 14 - Alg. 5): if not we remove b from $RN(x', y, wt')$. A similar step handles $LN(x', y, wt')$.

5.2 Description of FF-FIXUP

FF-FIXUP is an extension of the fixup in **PR** rather than **NPRdec**. This is because of the presence of the control bit β (defined in Section 4), and the need to process historical triples (that are not present in **NPRdec**). Algorithm FF-FIXUP will efficiently maintain exactly the LSTs and STs for each level graph in the tuple-system. This is in contrast to **PR**, which can maintain LHTs that are not LSTs in any level graph (PDG). FF-FIXUP maintains a heap H_f of candidate LHTs to be processed

in min-weight order. The main phase (Alg. 7) is very similar to the fixup in PR. The differences are again related to levels, centers and the new data structures.

We start describing Algorithm 7. We initialize H_f by inserting the edges incident on the updated vertex v with their updated weights (Steps 2 to 7 – Alg. 8), as well as a candidate min-weight triple from P for each pair of nodes (x, y) (Step 10 – Alg. 8). Then we process H_f by repeatedly extracting collections of triples of the same min-weight for a given pair of nodes, until H_f is empty (Steps 3 to 10 – Alg. 7). We will establish that the first set of triples for each pair (x, y) always represents the shortest path distance from x to y (see Lemma 7), and the triple extracted are added to the tuple-system if not already there (see Alg. 9 and Lemma 8). As in PR for efficiency, among all the triples present in the tuple-system for a pair of nodes, we select only the ones that need to be extended: this task is performed by Algorithm 9 (this step is explained later in the description). After the triples in S are left and right extended by Algorithm 10, we set the bit $\beta(\gamma') = 1$ for each triple γ' that is identified as shortest in S , since γ' is correctly updated both in $P^*(x, y)$ and $P(x, y)$ (Step 9 – Alg. 7). Finally, we update the $DL(x, y)$ structure by inserting (or updating if an element with weight wt is already present) the element with weight wt and the current level at the end of the list (Step 10 – Alg. 7). This concludes the description of Algorithm 7.

We now describe Algorithm 9 which is responsible to select only the triples that have valid extensions that will generate LHTs in the current graph. In Algorithm 9, we distinguish two cases. When the set of extracted triples from x to y contains at least one path not containing v (Step 2 – Alg. 9), then we process all the triples from $P(x, y)$ of the same weight. Otherwise, if all the paths extracted go through v (Step 18 – Alg. 9), we only use the triples extracted from H_f .

Algorithm 7 FF-FIXUP(v, \mathbf{w}', k)

```

1:  $H_f \leftarrow \emptyset$ ; Marked-Tuples  $\leftarrow \emptyset$ 
2: FF-POPULATE-HEAP( $v, \mathbf{w}', k$ )
3: while  $H_f \neq \emptyset$  do
4:   extract in  $S'$  all the triples with min-key  $[wt, x, y]$  from  $H_f$ 
5:   if  $S'$  is the first extracted set from  $H_f$  for  $x, y$  then
6:      $S \leftarrow$  FF-NEW-PATHS( $S', P(x, y)$ )
7:     FF-FIXUP- $\ell$ -extend( $S, [wt, x, y]$ ) (see Algorithm 10)
8:     FF-FIXUP- $r$ -extend( $S, [wt, x, y]$ )
9:     for every  $\gamma \in S$  set  $\beta(\gamma) = 1$ 
10:    add an element with weight  $wt$  and level  $k$  to  $DL(x, y)$  or update the level in the existing one

```

Algorithm 8 FF-POPULATE-HEAP(v, \mathbf{w}', k)

```

1: for each  $(u, v)$  do
2:    $\mathbf{w}(u, v) = \mathbf{w}'(u, v)$ 
3:   if  $\mathbf{w}(u, v) < \infty$  then
4:      $\gamma = ((uv, uv), \mathbf{w}(u, v), 1)$ ;  $C_\gamma[k] \leftarrow 1$ 
5:     update-num( $\gamma$ )  $\leftarrow$  curr-update-num; num-v-paths( $\gamma$ )  $\leftarrow 1$ 
6:     add  $[\gamma, C_\gamma]$  to  $H_f$  and  $P(u, v)$ 
7:     add  $u$  to  $L(-, vv)$  and  $v$  to  $R(uu, -)$ 
8: for each  $(v, u)$  do
9:   symmetric processing as Steps 2–7 above
10: for each  $x, y \in V$  do
11:   add a min-key triple  $[\gamma, C_\gamma] \in P(x, y)$  to  $H_f$ 

```

Both cases have a similar approach but here we focus on the former which is more involved than the latter. As soon as we identify a new triple γ' we compute its center j by using its associated array $C_{\gamma'}$ (Step 4 – Alg. 9). This is straightforward if compared to FF-CLEANUP where we first need

Algorithm 9 FF-NEW-PATHS(S', P_{xy})

```
1:  $S \leftarrow \emptyset$ ; let  $i$  be the min-weight level associated with  $DL(x, y)$ 
2: if  $P^*(x, y)$  increased min-weight after cleanup then
3:   for each  $\gamma' \in S$  with-key  $[wt, 0]$  do
4:     let  $\gamma' = ((xa', b'y), wt, count')$  and  $j = \text{argmin}_j(C_{\gamma'}[j] \neq 0)$ 
5:     if  $\gamma'$  is not in  $P^*(x, y)$  then
6:       add  $\gamma'$  in  $P^*(x, y)$  and  $S$ ; add  $x$  to  $L^*(a', y)$  and  $y$  to  $R^*(x, b')$ 
7:       add  $b'$  to  $RN(x, y, wt)$ ; place a double link between  $\gamma'$  and  $DL(x, y, wt)$ 
8:     else if  $\gamma'$  is in  $P(x, y)$  and  $P^*(x, y)$  with different counts then
9:       replace the count of  $\gamma'$  in  $P^*(x, y)$  with  $count'$  and add  $\gamma'$  to  $S$ 
10:      add  $\gamma'$  to  $P_j^*(x, y)$  and  $dict_j$ 
11:      add  $x$  to  $L_j^*(a', y)$  and  $y$  to  $R_j^*(x, b')$ 
12:      add  $x$  to  $LC_j^*(a', y)$  ( $y$  to  $RC_j^*(x, b')$ ) if  $x$  ( $y$ ) is a level  $i$  center
13:      add  $\gamma'$  in  $S$ 
14:   for each  $b' \in RN(x, y, wt)$  do
15:     if  $\exists h < i : L_h^*(x, b') \neq \emptyset$  then
16:       add any  $\gamma'$  of the form  $(x \times, b'y)$  and weight  $wt$  in  $P^*(x, y)$  with  $\beta(\gamma') = 1$  to  $S$ 
17: else
18:   for each  $\gamma' \in S'$  containing a path through  $v$  do
19:     let  $\gamma' = ((xa', b'y), wt, count')$  and  $k$  the current level
20:     add  $\gamma'$  with paths( $\gamma', v$ ) to  $P^*(x, y)$ , and  $[\gamma', C_{\gamma'}]$  to  $S$ 
21:     add  $\gamma'$  to  $P_k^*(x, y)$  and  $dict_k$ ,  $x$  to  $L_k^*(a', y)$  and  $y$  to  $R_k^*(x, b')$ 
22:     add  $x$  to  $LC_k^*(a', y)$  ( $y$  to  $RC_k^*(x, b')$ ) if  $x$  ( $y$ ) is a level  $k$  center
23: return  $S$ 
```

to update the center arrays. We add this triple to $P^*(x, y)$ and to S , which contains the set of triples that need to be extended. We also add γ' to $P_j^*(x, y)$ (Steps 10 and 21 – Alg. 9). We update $dict_j$ to keep track of the locations of the triple in the global structures. A similar sequence of steps takes place when all the extracted paths go through v (Steps 18 to 22 – Alg. 9). The only difference is that the local data structures to be updated are only the Γ_k data structures (Steps 21 and 22 – Alg. 9).

A crucial difference from PR and this algorithm is the way we collect the set S of triples to be extended. Here we require the new HE data structures RN and LN (see Section 4.2) because of PEP instances (see Section 6). Let i be the min-weight level associated with $DL(x, y)$. For each node $b \in RN(x, y, wt)$ we check if $L_h^*(x, b)$ contains at least one extension, for every $h < i$ (Steps 14 to 16 – Alg. 9). In fact we need to discover all tuples with $\beta = 1$ that are inside a PEP instance. In this instance, the triples restored as STs may or may not be extended. We cannot afford to look at all of them, thus our solution should check only the triples with an available extension. Moreover, all the extendable triples with $\beta = 1$ have extension only in levels younger than the level where they last appear as STs. Thus, we check for extensions only in the levels $h < i$.

Using the HE sets, is the key to avoid an otherwise long search of all the valid extensions for the set of examined triples with $\beta = 1$. In particular, without the HE sets, the algorithm could waste time by searching for extensions that are not even in the tuple-system. Correctness of this method is proven in section 7. After the algorithm collects the set S of triples that can be extended, FF-FIXUP calls FF-FIXUP- ℓ -extend (Alg. 10).

Here we describe the details of algorithm 10. Its goal is to generate LHTs for the current graph G by extending HTs. Let h be the *center* of S_b defined as the most recent center among all the triples in S_b , and let j be the level associated to the first weight wt' larger than wt in $DL(x, y)$. The extension phase for triples is different from PR: in fact, the set of triples S_b could contain only triples with $\beta(\gamma) = 1$. In PR, the corresponding set S_b contains only triples with $\beta(\gamma) = 0$. We address two cases:

(a) – If S_b contains at least one triple γ with $\beta(\gamma) = 0$, we extend S_b using the sets L_i^* and R_i^* with $h \leq i < j$ (Steps 7 to 16 – Alg.10). In fact, the set S_b contains at least one new path that was not

Algorithm 10 FF-FIXUP- ℓ -extend($S, [wt, x, y]$)

```

1: for every  $b$  such that  $(x \times, by) \in S$  do
2:   let  $S_b \subseteq S$  be the set of all triples of the form  $(x \times, by)$ 
3:   let  $fcount'$  be the sum of all the counts of all triples in  $S_b$ ; let  $h$  be the center( $S_b$ )
4:   if  $\exists \gamma \in S_b : \beta(\gamma) = 0$  then
5:     let  $j$  be the level associated to the minweight  $wt' > wt$  in  $DL(x, y)$ 
6:     for every active level  $h \leq i < j$  do
7:       for every  $x'$  in  $L_i^*(x, b)$  do
8:         if  $(x'x, by) \notin \text{Marked-Tuples}$  then
9:            $wt' \leftarrow wt + \mathbf{w}(x', x)$ ;  $\gamma' \leftarrow ((x'x, by), wt', fcount')$ 
10:           $C_{\gamma'} \leftarrow \text{FF-FIXUP-CENTERS}(S_b)$ ; add  $\gamma'$  to  $H_f$ 
11:          if a triple  $\gamma''$  for  $((x'x, by), wt')$  exists in  $P(x', y)$  then
12:            update the count of  $\gamma''$  in  $P(x', y)$  and  $C_{\gamma''} = C_{\gamma''} + C_{\gamma'}$ 
13:            add  $(x'x, by)$  to Marked-Tuples
14:          else
15:            add  $[\gamma', C_{\gamma'}]$  to  $P(x', y)$ ; add  $x'$  to  $L(x, by)$  and  $y$  to  $R(x'x, b)$ 
16:            set  $\beta(\gamma') = 0$ ; set  $\text{update-num}(\gamma')$ 
17:          for every level  $i < h$  do
18:            for every  $x'$  in  $LC_i^*(x, b)$  do
19:              execute steps 8 to 16
20:          else
21:            let  $j$  be the level associated to the minweight  $wt$  in  $DL(x, y)$ 
22:            for every level  $i < j$  do
23:              for every  $x'$  in  $LC_i^*(x, b)$  do
24:                execute steps 8 to 16

```

Algorithm 11 FF-FIXUP-CENTERS(S_b)

```

1: let  $C' = \sum_{\gamma \in S_b} C_\gamma$  be the sum (level by level) of the new paths that are found shortest
2: let  $j$  be  $\text{argmax}_j (C'[j] \neq 0)$ , and  $k = \text{center}(x')$ 
3: if  $k < j$  then
4:   for all the levels  $i < k$  we set  $C_{\gamma'}[i] = C'[i]$ 
5:   for the level  $k$  we set  $C_{\gamma'}[k] = \sum_{q=k}^{r-1} C'[q]$ 
6:   for all the levels  $m > k$  we set  $C_{\gamma'}[m] = 0$ 
7: else
8:    $C_{\gamma'} = C'$ 
9: return  $C_{\gamma'}$ 

```

extended in the previous iterations when wt was the shortest distance from x to y (because of the $\beta(\gamma) = 0$ triple). The LST generated in this way remains centered in level h . Moreover we extend S_b also using the sets LC_i^* and RC_i^* with $i < h$ (Steps 17 to 19 – Alg.10). This ensures that every LST generated in a level i lower than h is centered in i thanks to the extension node itself. This technique guarantees that each LHT generated by Algorithm 10 is an LST centered in a unique level.

(b) – In the case when there is no triple γ in S_b with $\beta(\gamma) = 0$, then there is at least one extension to perform for S_b and it must be in some level younger than the level where wt stopped to be the shortest distance from x to y (this follows from the use of the HE sets in Alg. 9). To perform these extensions we set j as the level associated with the min-weight element in $DL(x, y)$, and we extend S_b using the sets LC_i^* and RC_i^* with $i < j$ (Steps 21 to 24 – Alg.10). Again, every LHT generated is an LST centered in a unique level. Finally, following PR, every generated LHT is added to P and H_f and we update global L and R structures.

Observation 3 *Every LHT generated by algorithm FF-FIXUP is an LST centered in a unique level graph.*

Proof. As described in (a) and (b) above, every LHT is generated using two triples which are shortest in the same level graph Γ_i . Moreover, since at least one of them must be centered in level i , the resulting LHT is an LST centered in level i . \square

The last novelty in the algorithm is updating center arrays (Alg. 11 called at step 10 – Alg. 10) in a similar way of FF-CLEANUP: Algorithm 11 identifies the oldest level j related to the triples contained in S_b (Step 2 – Alg. 11). If $j > k$ then we reshape the distribution for γ' similarly to FF-CLEANUP (Steps 4 to 6 – Alg. 11). Otherwise γ' is completely contained in level k and no reshaping is required (Step 8 – Alg. 11).

6 New Features in Algorithm FFD

In this section, we discuss two challenges that arise when we attempt to generalize the level graph method used in Thorup (for APSP with unique SPs) to a fully dynamic APASP algorithm. Both are addressed by the algorithms in Section 5 as noted below.

The bit β feature. The control-bit β was introduced (and only briefly described) in PR to avoid the processing of untouched historical triples. Here, we elaborate on this technique in more detail than PR, and we also describe how it helps in the more complex setting of the level tuple-system.

Consider Figure 1. The ST $\gamma = ((xa, by), wt, count)$ is created in level k (Fig. 1(a)). At $time(k)$, we have $\gamma \in P^*$ and also $\gamma \in P$ with $\beta(\gamma) = 1$. In a more recent level $j < k$, a shorter triple $\gamma' = ((xv, vy), wt', count')$, with weight $wt' < wt$, that goes through an updated vertex v is generated (Fig. 1(b)). Thus at $time(j)$, we have $\gamma' \in P^*$ and also $\gamma' \in P$ with $\beta(\gamma') = 1$; but γ still appears in both P^* and P as a historical triple. Finally, a new LST $\gamma'' = ((xa', b'y), wt, count'')$, with the same weight as γ , is generated in level $i < j$ (Fig. 1(c)). Note that γ'' is only in P with $\beta(\gamma'') = 0$ and not in P^* , as is the case of every LST that is not an ST. When an increase-only update removes v and the triple γ' , the algorithm needs to restore all the triples with shortest weight wt . But while γ is historical and does not require any additional extension, γ'' is only present in P and needs to be processed. Our FFD algorithm achieves this by checking the bit β associated to each of these triples. The algorithm will extract and process all the triples with $\beta = 0$ from $P(x, y)$ (see Step 3, Alg. 9, where triples are extracted using $\beta = 0$ as part of the key). This guarantees that a triple only present in P , or present in P and P^* with different counts is never missed by the algorithm.

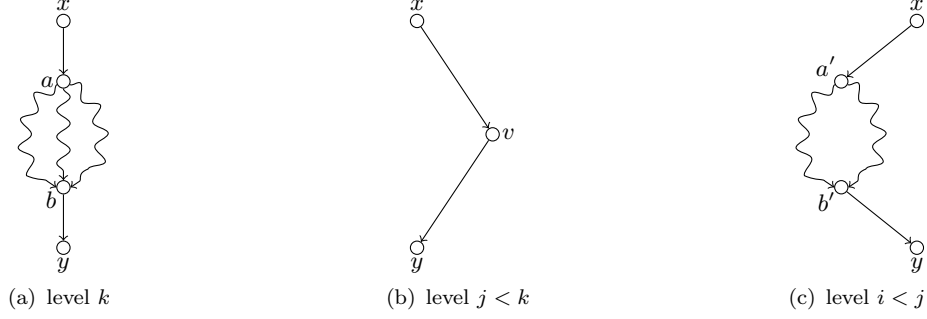


Fig. 1. The bit β feature

The partial extension problem (PEP). Consider the update sequence described below and illustrated Fig. 2. Here the STs $\gamma = ((xa, by), wt, count)$ and $\hat{\gamma} = ((xa, cy), wt, count')$ are created in level k (Fig. 2(a)). Later, a left-extension to x' generates the STs $\gamma' = ((x'a, by), wt', count)$ and $\hat{\gamma}' = ((x'a, cy), wt', count')$ in level $j < k$ (Fig. 2(b)). Note that γ , $\hat{\gamma}$, γ' and $\hat{\gamma}'$ are all present in P^* and P at $time(j)$. In a more recent level $i < j$, a decrease-only update on v generates a shorter triple $\gamma_s = ((xv, vy), wt_s, count_s)$ from x to y , with $wt_s < wt$ going through v . In the same level, the triple γ_s is also extended to x' generating a triple $\gamma'' = ((x'x, vy), wt'', count_s)$ shorter than γ' and $\hat{\gamma}'$ (Fig. 2(c)). Thus at $time(i)$, γ , $\hat{\gamma}$, γ' and $\hat{\gamma}'$ remain in P^* as historical triples. Then, in level $h < i$, an update on x'' inserts the edges (x'', x) and (x'', c) . This update generates an ST $\gamma''' = ((x''c, cy))$ (shorter than $(x''x, vy)$) and also inserts $x'' \in LC_h^*(x, b)$ since (x'', x) is on a shortest path from x'' to b ; but it should not generate the triple of the form $(x''x, by)$ because b is not on a shortest path from x to y at $time(h)$ (Fig. 2(d)).

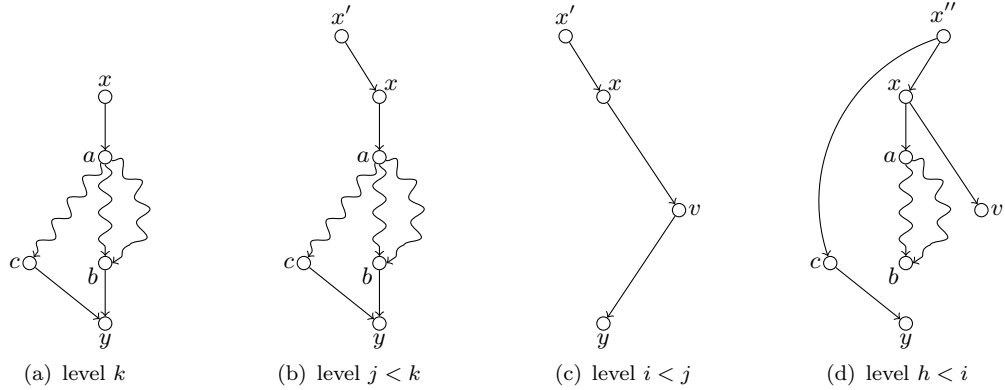


Fig. 2. PEP instance (only centered STs are kept in each level) – all edge weights are unitary

When an increase-only update removes v and the shortest triple γ_s from x to y , the algorithm needs to restore all historical triples with shortest weights from x to y . When γ and $\hat{\gamma}$ are restored, we need to perform suitable left extensions as follows. An extension to x'' is needed only for γ : in fact $\hat{\gamma}$ should not be extended to x'' because the ℓ -tuple of the newly generated tuple is not an ST in the graph. On the other hand, no extension to x' is needed since both γ' and $\hat{\gamma}'$ will be restored (from HT to ST). Our algorithm needs to distinguish all of these cases correctly and efficiently.

In order to maintain both correctness and efficiency in this scenario for APASP, we use two new data structures (described in Section 4.2): (1) the historical distance matrix DL that allow us to efficiently determine the most recent level graph in which an HT was an ST (see for example Steps 5 and 21, Alg. 10), and (2) the HE sets LN and RN that allow us to efficiently identify exactly those new extensions that need to be performed (see for example Steps 14 – 16, Alg. 9). The methodology of these data structures is fully discussed in the description of FF-FIXUP (Section 5.2). Note that, the PEP doesn't arise in **Thorup** because of the unique SP assumption: in fact when only a single SP of a given length is present in the graph for each pair of nodes, the algorithm can check for all the $O(n^2)$ paths maintained in each level and decide which one should be extended. Given the presence of multiple SPs in our setting, we cannot afford to look at each tuple in the tuple-system.

7 Correctness and Complexity

In this section we will first prove the complexity bounds of our FF-UPDATE algorithm, then we will establish correctness.

7.1 Complexity

The complexity analysis of algorithm FFD is similar to that for the PR algorithm. We highlight the following new elements:

1. As noted in Section 5.2 (see Observation 3), every triple created by FF-FIXUP is an LST in the level graph (PDG) in which is centered, and by the decremental only properties of level graphs, it will continue to be an LST in that level graph until it is removed. In contrast, PR can create LHTs by combining HTs not centered in any PDG. This results in an additional $\Theta(\log n)$ factor in the amortized bound there.
2. We can bound the number of LHTs that contain a given vertex u as $O(z' \cdot \nu^{*2})$, where z' is the number of active level graphs that contain vertex u and tuples passing through u (by Corollary 1). Given our level tuple-system, z' is clearly $O(\log n)$. In PR, this bound is $(z + z'^2)$ where z is the number of active PDGs, and z' is the number of PDGs that contain v .
3. We can show that the number of accesses to RN and LN , outside of the newly created tuples, is worst-case $O(n \cdot \nu^*)$ per call to FF-UPDATE. The overhead given by the level data structures is $O(\log n)$ for each access (see Lemma 5). These structures are not used in PR.

Lemma 1. *Let G be a graph after a sequence of calls to FF-UPDATE. Let z be the number of active level graphs (PDGs), and let $z' \leq z$ be the number of level graphs that contain a given vertex v . Suppose that every HT in the tuple-system is an ST in some level graphs, and every LHT is an LST in some level graph. If n and m bound the number of vertices and edges, respectively, in any of these graphs, and if ν^* bounds the maximum number of distinct edges that lie on shortest paths through any given vertex in any of these graphs, then:*

1. *The number of LHTs in G 's tuple-system is at most $O(z \cdot m \cdot \nu^*)$.*
2. *The number of LHTs that contain a vertex v in G is $O(z' \cdot \nu^{*2})$.*

Proof. For part 1, we bound the number of LHTs (xa, by) (across all weights) that can exist in G . The edge (x, a) can be chosen in m ways, and once we fix (x, a) , the r -tuple (a, by) must be an ST in one of the Γ_j . Since (b, y) must lie on a shortest path through a centered in a graph Γ_i , that contains the r -tuple (a, by) of shortest weight in Γ_i , the number of different choices for (b, y) that

will then uniquely determine the tuple (xa, by) , together with its weight, is $z \cdot \nu^*$. Hence the number of LHTs in G 's tuple-system is $O(z \cdot m \cdot \nu^*)$.

For part 2, the number of LHTs that contain v as an internal vertex is simply the number of LSTs across the z' graphs that contains v , and this is $O(z' \cdot \nu^{*2})$. We now bound the number of LHTs (va, by) . There are $n - 1$ choices for the edge (v, a) and $z' \cdot \nu^*$ choices for the r -tuple (a, by) , hence the total number of such tuples is $O(z' \cdot n \cdot \nu^*)$. The same bound holds for LHTs of the form (xa, bv) . Since $\nu^* = \Omega(n)$, the result in part 2 follows. \square

Corollary 1. *At a given time step, let B be the maximum number of tuples in the tuple-system containing a path through a given vertex in a given level graph. Then, $B = O(\nu^{*2})$.*

Lemma 2. (a) - *The cost for an FF-CLEANUP call on a node v when z' active levels contain triples through v is $O(z' \cdot \nu^{*2} \cdot \log n)$.*

(b) - *The cost for a real FF-CLEANUP call is $O(\nu^{*2} \cdot \log^2 n)$*

(c) - *The cost for a dummy FF-CLEANUP call is $O(\nu^{*2} \cdot \log n)$.*

Proof. (a) - Algorithm 4 extracts all the LHTs that go through the update vertex from the heap H_c . Since the number of these LHTs is bounded by B at each level (by Corollary 1), the total cost is $O(z' \cdot B \log n)$ where z' is the number of active levels that contain triples through v . Algorithm 6 requires only $O(\log n)$ time for each step, except in step 4 where the cost is $O(\log n)$ for each triple extracted from H_c that goes through the updated vertex. Since the number of such triples is bounded by $O(z' \cdot \nu^{*2})$ (by Lemma 1), the worst-case cost for a call to Algorithm 6 within an FF-CLEANUP phase is $O(z' \cdot \nu^{*2} \cdot \log n)$. In Algorithm 5, a triple can be added to heap H_c , or searched and removed from a constant number of priority queues among z' different active levels. Moreover, for the structures DL , RN and LN each triple spends a constant time to be unlinked and eventually to update the structures. Since, priority queue operations have a $O(\log n)$ cost and the number of triples examined is bounded by $O(z' \cdot \nu^{*2})$, the complexity for Algorithm 5 is at most $O(z' \cdot \nu^{*2} \cdot \log n)$. Thus an FF-CLEANUP call that operates on z' active levels requires at most $O(z' \cdot \nu^{*2} \cdot \log n)$.

(b) - Since $z \leq \log 2n$, the cost for a real FF-CLEANUP call is $O(\nu^{*2} \cdot \log^2 n)$ (by part (a)).

(c) - For a dummy cleanup on a vertex w , FF-CLEANUP only needs to clean the local data structures in level $center(w)$, where w is centered, and in the current level graph. In fact, let t be the current update step; in the dummy cleanup phase, we start with the node u that was updated at time $t - 1$ (the most recent update before the current one). The node u received an update in the previous phase, thus it disappeared from all the levels older than $level(t - 1)$ and, with it, all the LSTs containing u in these levels. Hence, all the triples containing u in the tuple-system must be LSTs in $level(t - 1)$. We have at most B of them and FF-CLEANUP spends $O(B \cdot \log n)$ (considering the access to the data structures) to remove them. Then, the dummy update reinserts u only in the current graph. The next phase moves on the node u' updated at time $t - 2$. Again, all the tuples containing u' must be LSTs in $level(t - 2)$ and eventually the current graph if they were inserted because of the previous dummy update on u .

Suppose in fact that there is a tuple γ that contains u' in another level (except the current graph). The tuple γ cannot be in a level older than $level(t - 2)$ because when u' was updated at time $(t - 2)$, the cleanup algorithm removed all the tuples containing u' from any level older than $t - 2$. Moreover, a tuple containing u' present in a level younger than $level(t - 2)$ could appear if and only if it was generated by any update more recent of $t - 2$ (in this case only the dummy update on u performed in the current graph). Thus a contradiction.

This argument can be recursively applied to every other node in the sequence: in fact for the node u'' updated at time $(t - i)$ all the nodes updated in the interval $[t - i + 1, t - 1]$ will be already

processed by FF-CLEANUP, leaving all the tuples containing u'' only in $level(t - i)$ and t . It follows that, for a dummy update, $z' = 2$. Thus the cost for a dummy FF-CLEANUP call is $O(\nu^{*2} \cdot \log n)$ (by part (a)). \square

Lemma 3. *The cost for a dummy FF-FIXUP call on a node v is $O(\nu^{*2} \cdot \log n)$.*

Proof. Consider a dummy FF-FIXUP applied to a vertex v in \mathcal{N} . We only need to bound the cost for accessing the entries in the $P^*(x, y)$ and the cost of re-adding LSTs containing v , previously removed by the dummy FF-CLEANUP but still in the current graph after the dummy update. In fact the vertex v is removed by an earlier dummy FF-CLEANUP, and while this removes all the HPs containing the vertex v , it does not change any LST centered in any Γ_j that does not contain v . Hence these other LSTs will be present in the tuple-system with unchanged weight and count, when dummy FF-FIXUP is applied to v . Since for any pair x, y , the SP distance will not change after the dummy update, the dummy FF-FIXUP will only insert in the set S triples containing the node v for additional extension (this is accomplished by the check at Step 2 - Alg. 9, followed by Steps 18–22, Alg. 9). Hence, only the LSTs containing v in the current $level(t)$ graph will be processed and added to the tuple-system, and there are at most B of them (by Corollary 1). Thus a dummy FF-FIXUP for any v needs to access P^* for each pair of nodes, and reinsert at most B tuples (containing v) in the current graph. Hence the overall complexity for a dummy FF-FIXUP is $O((n^2 + B) \cdot \log n) = O(\nu^{*2} \cdot \log n)$. \square

We now address the complexity of a real FF-FIXUP call. We first define the concept of a *triple pair* that will be used in lemma 5 to establish the bound for a real FF-FIXUP call. Finally, we complete our analysis by presenting a proof of Theorem 1.

Definition 1. *If $C_\gamma[i] \geq 1$ then (γ, i) is a triple pair in the tuple-system. If (γ, i) is not a triple pair in the tuple-system at the start of step t but is a triple pair after the update at time step t , then (γ, i) is a newly created triple pair at time step t .*

Lemma 4. *At a given time step, let D be the number of triple pairs in the level tuple-system. Then,*

1. *The value of D is at most $O(m \cdot \nu^* \cdot \log n)$.*
2. *The space used is $O(m \cdot \nu^* \cdot \log n)$.*

Proof. 1. Every $C_\gamma[i] \geq 1$ represents a distinct LST in Γ_i , hence the result follows since the number of levels is $O(\log n)$ and the number of LSTs in a graph is $O(\nu^* \cdot m^*)$.

2. Since every triple is of size $O(1)$, the memory used by our FFD algorithm is dominated by D , and result follows from 1. \square

Lemma 5. *The cost for a real FF-FIXUP call is $O(\nu^{*2} \cdot \log^2 n + X \cdot \log n)$, where X is the number of newly created triple pairs after the update step.*

Proof. A triple is accessed only a constant number of time during FF-FIXUP for a cost of $O(\log n)$, so it suffices to establish that the number of existing triples accessed during the call is $O(\nu^{*2} \cdot \log n)$.

There are only $O(n^2)$ accesses to triples in the call to Algorithm 8 in line 2 of FF-FIXUP since $O(n^2)$ entries in the global $P^*(x, y)$ structures are accessed to initialize H_f . This takes $O(n^2 \cdot \log n)$ time after considering the $O(\log n)$ cost per data structure operation. We now address the accesses made in the main loop. We will distinguish two cases and they will be charged to X as follows.

1: $\beta(\gamma) = 0$ – In the main loop of Algorithm 7, starting in Step 3, any triple γ that is accessed with $\beta(\gamma) = 0$ is an LST at some level i where it is not identified as an ST in Γ_i . During this call, in

Steps 3–12, Alg. 9 (or Steps 18–22, Alg. 9 if the distance for the endpoints of γ did not change) γ is added as an ST in level i , and will never be removed as an ST for level i until it is removed from the tuple system (due to the fact that T_i is a purely decremental graph). Since γ with $\beta(\gamma) = 0$ is a newly added triple to level i , then the pair (γ, i) is a newly created triple pair at step t . Hence, we can charge (γ, i) to X in this call of FF-FIXUP.

2: $\beta(\gamma) = 1$ – We now consider triples accessed that have $\beta = 1$. This is the most nontrivial part of our analysis since even though any such triple γ must exist with the same count in every level in both P and P^* , we may still need to form some extensions since the triple may have been an HT when extension vertices were updated, and hence these extension may not have been performed. Here is where the LN and RN sets are accessed, and we now analyze the cost of these accesses. (The correctness of the associated steps is analyzed in the next section.)

Let j be the most recent level in which γ was an ST in G and assume we are dealing with left extensions (right extensions are symmetrical). Now that γ is restored, the only case (Steps 14 to 16, Alg. 9) in which we need to process it is when there exists a left extension for the ℓ -tuple of γ to a node x' centered in a level i more recent than j . In fact, the LST generated by this extension will appear for the first time centered in level i , hence the pair (γ, i) is a newly created triple pair at step t and we can charge its creation to X . We now show how our HE sets efficiently handle this case. By steps 14 to 16, Alg. 9, our algorithm only processes a restored triple γ with $\beta(\gamma) = 1$ when it has at least one centered extension in some active level younger than the level in which γ was shortest for the last time. We can bound the total computation for these steps as follows: for a given x , $RN(x, y, wt)$ contains a node b for every incoming edge to y in one of the SSSP dags (historical and shortest) rooted at x . Since we can extend in at most $O(\log n)$ active levels during any update and the size of a single dag is at most ν^* , these steps take time $O(\nu^* \cdot n \log n)$ throughout the entire update computation. \square

We can now establish the proof of our main theorem.

Proof of Theorem 1. Consider a sequence Σ of $r = \Omega(n)$ calls to algorithm FFD. Recall that the data structure is reconstructed after every $2n$ steps, so we can assume $r = \Theta(n)$. These r calls to FFD make r real calls to FF-UPDATE, and also make additional dummy updates. As in PR, across the r real updates in Σ , the algorithm performs $O(r \log n)$ dummy updates. This is because $r/2^k$ real updates are performed at level k during the entire computation, and each such update is accompanied by $2^k - 1$ dummy updates. So, across all real updates there are $O(r)$ dummy updates per level, adding up to $O(r \log n)$ in total, across the $O(\log n)$ levels.

When FF-CLEANUP is called on a vertex v for a dummy update, $z' = 2$ since v can be present only in the most recent current level and the level at which it is centered. (This is because every vertex that was centered at a more recent level than v has already been subjected to a dummy update, and hence all of these vertices are now centered in the current level.) Thus, by Lemma 2, each FF-CLEANUP for a dummy update has cost $O(B \cdot \log n)$. By Lemma 3, a call to FF-FIXUP for a dummy update has cost $O(\nu^{*2} \cdot \log n)$. Thus the total cost is $O((\nu^{*2} \cdot \log n) \cdot r \log n)$ across all dummy updates. Also, the number of tuples accessed by all of the dummy update calls to FF-CLEANUP, and hence the number of tuples removed by all dummy updates, is $O(r \cdot \nu^{*2} \cdot \log n)$.

For the real calls to FF-FIXUP, let X_i be the number of newly added triple pairs in the i th real call to FF-FIXUP. Then by Lemma 5, the cost of this i th call is $O(\nu^{*2} \cdot \log^2 n + X_i \cdot \log n)$. Let $X = \sum_{i=1}^r X_i$. Hence the total cost for the r real calls to FF-FIXUP is $O(r \cdot \nu^{*2} \cdot \log^2 n + X \cdot \log n)$. We now bound X as follows: X is no more than the maximum number of triples that can remain in the system after Σ is executed, plus the number of tuples Y removed from the tuple-system. Tuples are removed only in calls to FF-CLEANUP. The total number removed by $r \log n$ dummy calls is $O(r \cdot \log n \cdot \nu^{*2})$ (by Lemma 2). The total number removed by the r real calls is $O(r \cdot \nu^{*2} \cdot \log n)$ (by Lemma 2). Hence

$Y = O(r \cdot \nu^{*2} \cdot \log n)$. Clearly the maximum number of triples in the tuple-system is no more than D , which counts the number of triple pairs, and we have $D = O(m \cdot \nu^* \cdot \log n) = O(n^2 \cdot \nu^* \cdot \log n)$ (by Lemma 4). Since $r = \Theta(n)$, we have $D = O(r \cdot n \cdot \nu^* \cdot \log n)$, and this is dominated by Y since $\nu^* = \Omega(n)$. Hence the cost of the r calls to FFD is $O(r \cdot \nu^{*2} \cdot \log^2 n)$ (after factoring in the $O(\log n)$ cost per tuple access), and hence the amortized cost of each call to FF-UPDATE is $O(\nu^{*2} \cdot \log^2 n)$.

7.2 Correctness

For the correctness, we assume that all the global and local data structures are correct before the update, and we will show the correctness of them after the update.

Correctness of Cleanup - The correctness of FF-CLEANUP is established in Lemma 6. We will prove that all paths containing the updated vertex v are removed from the tuple-system. Moreover, the center of each triple is restored, if necessary, to the level containing the most recently updated node on any path in this triple. Note that (as in [8,19,21]) at the end of the cleanup phase, the global structures P and P^* may not have all the LHTs in $G \setminus \{v\}$.

Lemma 6. *At the end of the cleanup phase triggered by an update on a vertex v , every LHP that goes through v is removed from the global structures. Moreover, in each level graph Γ_i , each SP that goes through v is removed from P_i^* . For each level i , the local structures L_i^* , R_i^* , RC_i^* and LC_i^* contain the correct extensions; the global structures L and R contain the correct extensions, for each r -tuple and ℓ -tuple respectively, and the structures RN and LN contain only nodes associated with tuples in P . The DL structure only contains historical distances represented by at least one path in the updated graph. Finally, every triple in P and P^* has the correct updated center for the graph $G \setminus \{v\}$.*

Proof. To prove the lemma statement, we use a loop invariant on the while loop in Step 3 of Algorithm 4. We show that the while loop maintains the following invariants.

Loop Invariant: At the start of each iteration of the while loop in Step 3 of Algorithm 4, assume that the first triple to be extracted from H_c and processed has min-key = $[wt, x, y]$. Then the following properties hold about the tuple-system and H_c .

1. For any $a, b \in V$, if G contains c_{ab} LHPs of weight wt of the form (xa, by) passing through v , then H_c contains a triple $\gamma = ((xa, by), wt, c_{ab})$ with key $[wt, x, y]$ already processed: the c_{ab} LHPs through v are not present in the tuple-system.
2. Let $[\hat{wt}, \hat{x}, \hat{y}]$ be the last key extracted from H_c and processed before $[wt, x, y]$. For any key $[wt_1, x_1, y_1] \leq [\hat{wt}, \hat{x}, \hat{y}]$, let G contain $c > 0$ number of LHPs of weight wt_1 of the form $(x_1 \times, b_1 y_1)$. Further, let c_v (resp. $c_{\bar{v}}$) denote the number of such LHPs that pass through v (resp. do not pass through v). Here $c_v + c_{\bar{v}} = c$. For every extension $x' \in L(x_1, b_1 y_1)$, let $wt' = wt_1 + \mathbf{w}(x', x_1)$ be the weight of the extended triple $(x' x_1, b_1 y_1)$. Then, (the following assertions are similar for $y' \in R(x_1 a_1, y_1)$)

Global Data Structures:

- (a) if $c > c_v$ there is a triple in $P(x', y_1)$ of the form $(x' x_1, b_1 y_1)$ and weight wt' representing $c - c_v$ LHPs. Moreover, its center is updated according to the last update on any path represented by the triple. If $c = c_v$ there is no such triple in $P(x', y_1)$.
- (b) If a triple of the form $(x' x_1, b_1 y_1)$ and weight wt' is present as an HT in $P^*(x', y_1)$, then it represents the exact same number of LHPs $c - c_v$ of the corresponding triple in $P(x', y_1)$. This is exactly the number of HPs of the form $(x' x_1, b_1 y_1)$ and weight wt' in $G \setminus \{v\}$.
- (c) $x' \in L(x_1, b_1 y_1)$, $y_1 \in R(x' x_1, b_1)$, and $(x' x_1, b_1 y_1) \in \text{Marked-Tuples}$ iff $c_{\bar{v}} > 0$.

- (d) A triple corresponding to $(x'x_1, b_1y_1)$ with weight wt' and counts c_v is in H_c . A similar assertion holds for $y' \in R(x_1a_1, y_1)$.
- (e) The structure $RN(x', y_1, wt')$ contains a node b iff at least one path of the form $(x' \times, by_1)$ and weight wt' is still represented by a triple in $P(x', y_1)$. A similar assertion holds for a node a in $LN(x', y_1, wt')$.
- (f) If there is no HT of the form $(x'x, b_1y_1)$ and weight wt' in $P^*(x', y_1)$ then the entry $DL(x', y_1)$ with weight wt' does not exist.

Local Data Structures: for each level j , let c_j be the number of LSPs of the form $(x'x_1, b_1y_1)$ and weight wt' centered in Γ_j and let $c_j(v)$ be the ones that go through v . Thus $c = \sum_j c_j$ and $c_v = \sum_j c_j(v)$. Then,

- (g) the value of $C_\gamma[j]$, where γ is the triple of the form $(x'x_1, b_1y_1)$ and weight wt' in $P(x', y_1)$, is $c_j - c_j(v)$.
 - (h) If a triple γ of the form $(x'x_1, b_1y_1)$ and weight wt' is present as an HT in P^* , then $P_j^*(x', y_1)$ represents only $c_j - c_j(v)$ paths. If $c_j - c_j(v) = 0$ then the link to γ is removed from $dict_j$. Moreover, $x' \in L_j^*(x_1, y_1)$ (respectively $LC_j^*(x_1, y_1)$ if x' is centered in Γ_j) iff x' is part of a shortest path of the form $(x'x_1, \times y_1)$ centered in Γ_j . A similar statement holds for $y_1 \in R_j^*(x', b_1)$ (respectively $RC_j^*(x', b_1)$ if y_1 is centered in Γ_j).
3. For any key $[wt_2, x_2, y_2] \geq [wt, x, y]$, let G contain $c > 0$ LHPs of weight wt_2 of the form (x_2a_2, b_2y_2) . Further, let c_v (resp. $c_{\bar{v}}$) denote the number of such LHPs that pass through v (resp. do not pass through v). Here $c_v + c_{\bar{v}} = c$. Then the tuple $(x_2a_2, b_2y_2) \in \text{Marked-Tuples}$, iff $c_{\bar{v}} > 0$ and a triple for (x_2a_2, b_2y_2) is present in H_c

Initialization: We start by showing that the invariants hold before the first loop iteration. The min-key triple in H_c has key $[0, v, v]$. Invariant assertion 1 holds since we inserted into H_c the trivial triple of weight 0 corresponding to the vertex v and that is the only triple of such key. Moreover, since we do not represent trivial paths containing the single vertex, no counts need to be decremented. Since we assume positive edge weights, there are no LHPs in G of weight less than zero. Thus all the points of invariant assertion 2 hold trivially. Invariant assertion 3 holds since H_c does not contain any triple of weight > 0 and we initialized Marked-Tuples to empty.

Maintenance: Assume that the invariants are true before an iteration k of the loop. We prove that the invariant assertions remain true before the next iteration $k + 1$. Let the min-key triple at the beginning of the k -th iteration be $[wt_k, x_k, y_k]$. By invariant assertion 1, we know that for any a_i, b_j , if there exists a triple γ of the form (x_ka_i, b_jy_k) of weight wt_k representing *count* paths containing v , then it is present in H_c . Now consider the set of triples with key $[wt_k, x_k, y_k]$ which we extract in the set S (Step 4, Algorithm 4). We consider left-extensions of triples in S ; symmetric arguments apply for right-extensions. Consider for a particular b the set $S_b \subseteq S$ of triples of the form (x_k-, by_k) , and let $fcount'$ denote the sum of the counts of the paths represented by triples in S_b . Let $x' \in L(x_k, by_k)$ be a left extension; our goal is to generate the triple γ' of the form $(x'x_k, by_k)$ with count $fcount'$ and weight $wt' = wt_k + \mathbf{w}(x', x_k)$, and an associated vector $C(\gamma')$ that specifies the distribution of paths represented by γ' level by level. These paths will be then removed by the algorithm. However, we generate such triple only if it has not been generated by a right-extension of another set of paths by checking the Marked-Tuples structure: we observe that the paths of the form $(x'x_k, by_k)$ can be generated by right extending to y_k the set of triples of the form $(x'x_k, \times b)$. Without loss of generality assume that the triples of the form $(x'x_k, \times b)$ have a key which is greater than the key $[wt_k, x_k, y_k]$. Thus, at the beginning of the k -th iteration, by invariant assertion 3, we know that $(x'x_k, by_k) \notin \text{Marked-Tuples}$. Step 5, Alg. 4 creates a triple γ' of the form $(x'x_k, by_k)$ of weight wt' and $fcount'$.

The set of triples in S_b can have different centers and we are going to remove (level by level) paths represented by γ' . To perform this task we consider the vector $C_{\gamma''}$: it contains the full

distribution of the triple $\gamma'' \in P(x'y_k)$ of the form $((x'x_k, by_k), wt')$ and indicates the oldest level j in which γ'' was generated for the first time. This level is exactly $\text{argmax}_j(C_{\gamma''}[j] \neq 0)$ and it is identified in Step 3 - Alg. 5. All the paths represented in S_b that are centered in some level m older or equal to j were extended for the first time in level j to generate γ'' . Moreover, each path centered in a level i younger than j was extended in level i itself. Thus, we can compute a new center vector $C_{\gamma'}$ (according to the distribution in $C_{\gamma''}$) of the paths containing v that we want to delete at each active level, as in steps 5 to 8 - Alg. 6. In step 9 - Alg. 6 the vector $C_{\gamma''}$ is updated: the paths are removed level by level according to the new distribution. This establishes invariant assertion 2g.

The triple γ' is immediately added to H_c with $C_{\gamma'}$ for further extensions (Step. 7 - Alg. 5). This establishes invariant assertions 2d. Thus we reduce the counts of γ' in $P(x', y_k)$ by $fcount$ (Step. 8 - Alg.5) and we set the new center for the remaining tuple γ'' in $P(x', y_k)$ establishing invariant assertion 2a. Steps 10 to 13 - Alg. 5 check if there is any path of the form $(x-, by)$ that can use x' as an extension. In this case we add γ' to the Marked-Tuples. If not, we safely remove the left and right extension (x' and y) from the tuple-system. This establishes invariant assertion 2c. If γ' is an HT in $P^*(x', y_k)$, we decrement its count (Step. 19 - Alg.5) establishing invariant assertion 2b. In steps 14 to 17 - Alg. 5, we use the double links between $b \in RN(x', y_k, wt')$ and tuples to efficiently check if there are other triples linked to b ; if not we remove b from $RN(x', y_k, wt')$ establishing invariant assertion 2e. Using a similar double link method with the structure $DL(x', y_k)$, we establish invariant assertion 2f after step 21 - Alg. 5.

To operate in the local data structures we require γ' to be an HT in $P^*(x', y_k)$. Using the previously created vector $C_{\gamma'}$, we reduce the count associated with $\gamma' \in 1P_i^*(x', y_k)$ for each level i (Step 23 - Alg. 5). After the above step, if there are no paths left in $P_i^*(x', y_1)$ then there are no STs of the form $(x'x_1, by_1)$ centered in level i . In this case we remove the extension x' and y_k from the local structures of level i . This is done in steps 24 to 34 -Alg. 5: in case x' is not centered in level i , then any path in γ' centered in level i is generated by a node centered in level i located between x_k and y_k . Thus if any SP from x' to y_k (that uses (x', x_k) as a first edge) remains in Γ_i , it must be also counted in $P_i^*(x_k, y_k)$. Thus, we remove x' from $L_i^*(x_k, y_k)$ only if $P_i^*(x_k, y_k)$ is empty. In the case x' is centered in level i and $P_i^*(x_k, y_k)$ is empty, x' could still be the extension of other paths from x_k to y_k centered in levels older than i . The algorithm checks them all and if they do not exist in older levels we can safely remove x' from $LC_i^*(x_k, y_k)$ (Step 32, Alg. 5). A similar argument holds for the right extension y_1 . This establishes invariant assertion 2h and completes claim 2.

When any triple is generated by a left extension (or symmetrically right extension), it is inserted into H_c as well as into Marked-Tuples. This establishes invariant assertion 3 at the beginning of the $(k+1)$ -th iteration.

Finally, to see that invariant assertion 1 holds at the beginning of the $(k+1)$ -th iteration, let the min-key at the $(k+1)$ -th iteration be $[wt_{k+1}, x_{k+1}, y_{k+1}]$. Observe that triples with weight wt_{k+1} starting with x_{k+1} and ending in y_{k+1} can be created either by left extending or right extending the triples of smaller weight. And since for each of iteration $\leq k$, invariant assertion 2 holds for any extension, we conclude that invariant assertion 1 holds at the beginning of the $(k+1)$ -th iteration. This concludes our maintenance step.

Termination: The condition to exit the loop is $H_c = \emptyset$. Because invariant assertion 1 maintains in H_c all the triples already processed, then $H_c = \emptyset$ implies that there are no other triples to extend in the graph G that contain the updated node v . Moreover, because of invariant assertion 1, every triple containing the node v inserted into H_c has been correctly decremented from the tuple-system. Remaining triples have the correct update center because of invariant 2a. Finally, for invariant assertions 2g and 2h, the structures $L_i^*, LC_i^*, R_i^*, RC_i^*$ are correctly maintained for every

active level i and the paths are surgically removed only from the levels in which they are centered. This completes the proof. \square

Correctness of Fixup - For the fixup phase, we need to show that the triples generated by our algorithm are sufficient to maintain all the ST and LST in the current graph G . As in PR, we first show in the following lemma that FF-FIXUP computes all the correct distances for each pair of nodes in the updated graph. Finally, we show that data structures and counts are correctly maintained at the end of the algorithm (Lemma 8).

Lemma 7. *For every pair of nodes (x, y) , let $\gamma = ((xa, by), wt, count)$ be one of the min-weight triples from x to y extracted from H_f during FF-FIXUP. Then wt is the shortest path distance from x to y in G after the update.*

Proof. Suppose that the lemma is violated. Thus, there will be an extraction from H_f during FF-FIXUP such that the set of extracted triples S' , of weight \hat{wt} is not shortest in G after the update. Consider the earliest of these events when S' is extracted from H_f . Since S' is not a set of STs from x to y , there is at least one shorter tuple from x and y in the updated graph. Let $\gamma' = ((xa', b'y), wt, count)$ be this triple that represents at least one shortest path from x to y , with $wt < \hat{wt}$. Since S' is extracted from H_f before any other triple from x to y , γ' cannot be in H_f at any time during FF-FIXUP. Hence, it is also not present in $P(x, y)$ as an LST at the beginning of the algorithm, otherwise it (or another triple with the same weight) would be placed in H_f by step 2 - Alg. 7. Moreover, if γ' is a single edge (trivial triple), then it was already an LST in G present in $P(x, y)$ before the update, and it is added to H_f by step 10 - Alg. 8; moreover since all the edges incident to v are added to H_f during steps 2 to 7 of Alg. 8, then γ' must represent SPs of at least two edges. We define $left(\gamma')$ as the set of LSTs of the form $((xa', c_i b'), wt - \mathbf{w}(b', y), count_{c_i})$ that represent all the LSPs in the left tuple $((xa', b'), wt - \mathbf{w}(b', y))$; similarly we define $right(\gamma')$ as the set of LSTs of the form $((a' d_j, b'y), wt - \mathbf{w}(x, a'), count_{d_j})$ that represent all the LSPs in the right tuple $((a', b'y), wt - \mathbf{w}(x, a'))$.

Observe that since γ' is an ST, all the LSTs in $left(\gamma')$ and $right(\gamma')$ are also STs. A triple in $left(\gamma')$ and a triple in $right(\gamma')$ cannot be present in P^* together at the beginning of FF-FIXUP. In fact, if at least one triple from both sets is present in P^* at the beginning of FF-FIXUP, then the last one inserted during the fixup phase triggered during the previous update, would have generated an LST of the form $((xa', b'y), wt)$ automatically inserted, and thus present, in P at the beginning of the current fixup phase (a contradiction). Thus either there is no triple represented by $left(\gamma')$ in P^* , or there is no triple represented by $right(\gamma')$ in P^* .

Assume w.l.o.g. that the set of triples in $right(\gamma')$ is placed into P^* after $left(\gamma')$ by FF-FIXUP. Since edge weights are positive, $wt - \mathbf{w}(x, a') < wt < \hat{wt}$, and because all the extractions before γ were correct, then the triples in $right(\gamma')$ were correctly extracted from H_f and placed in P^* before the wrong extraction of S' . Let i be the level in which $left(\gamma')$ is centered, and let j be the level in which $right(\gamma')$ is centered. By the assumptions, all the triples in $left(\gamma')$ are in P^* and we need to distinguish 3 cases:

1. if $j = i$, then FF-FIXUP generates the tuple $((xa', b'y), wt)$ in the same level and place it in P and H_f .
2. if $i > j$, the algorithms FF-FIXUP extends the set $right(\gamma')$ to all nodes in $L_i^*(a', b')$ for every $i \geq j$ (see Steps 7 to 16 - Alg. 10). Thus, since $left(\gamma')$ is centered in some level $i > j$, the node x is a valid extension in $L_i^*(a', b')$, making the generated γ' an LST in Γ_j that will be placed in $P(x, y)$ and also into H_f (during Step 10 - Alg. 10).

3. if $j > i$, then x was inserted in a level younger than i . In fact, all the paths from a' to b' must be the same in $right(\gamma')$ and $left(\gamma')$ otherwise the center of $right(\gamma')$ should be i . Hence, the only case when $j > i$ is when the last update on $left(\gamma')$ is on the node x in a level i younger than j . Thus $x \in LC_i^*(a', b')$. But FF-FIXUP extends $right(\gamma')$ to all nodes in $LC_i^*(a', b')$ for every $i < j$, placing the generated LST γ' in $P(x, y)$ and also into H_f (see Steps 17 to 19 - Alg. 10).

Thus the algorithm would generate the tuple $((xa', b'y), wt)$ (as a left extension) and place it in P and H_f (because all the triples in $left(\gamma')$ are already in P^*). Therefore, in all cases, a tuple $((xa', b'y), wt)$ should have been extracted from H_f before any triple in S' . A contradiction. \square

Lemma 8. *After the execution of FF-FIXUP, for any $(x, y) \in V$, the sets $P^*(x, y)$ ($P(x, y)$) contains all the SPs (LSPs) from x to y in the updated graph. Also, the global structures L, R and the local structures $P_i^*, L_i^*, R_i^*, LC_i^*, RC_i^*$ and $dict_i$ for each level i are correctly maintained. The structures RN and LN are updated according to the newly identified tuples. The DL structure contains the updated distance for each pair of nodes in the current graph. Finally, the center of each new triple is updated.*

Proof. We prove the lemma statement by showing the following loop invariant. Let G' be the graph after the update.

Loop Invariant: At the start of each iteration of the while loop in Step 3 of FF-FIXUP, assume that the first triple in H_f to be extracted and processed has $\text{min-key} = [wt, x, y]$. Then the following properties hold about the tuple-system and H_f .

1. For any $a, b \in V$, if G' contains c_{ab} SPs of form (xa, by) and weight wt , then H_f contains a triple of form (xa, by) and weight wt to be extracted and processed. Further, a triple $\gamma = ((xa, by), wt, c_{ab})$ is present in $P(x, y)$.
2. Let $[\hat{wt}, \hat{x}, \hat{y}]$ be the last key extracted from H_f and processed before $[wt, x, y]$. For any key $[wt_1, x_1, y_1] \leq [\hat{wt}, \hat{x}, \hat{y}]$, let G' contain $c > 0$ number of LHPs of weight wt_1 of the form (x_1a_1, b_1y_1) . Further, let c_{new} (resp. c_{old}) denote the number of these LHPs that are *new* (resp. not *new*). Here $c_{new} + c_{old} = c$. If $c_{new} > 0$ then,

Global Data Structures:

- (a) there is an LHT γ in $P(x_1, y_1)$ of the form (x_1a_1, b_1y_1) and weight wt_1 that represents c LHPs, with an updated center defined by the last update on any of the paths represented by the LHT.
- (b) If a triple of the form (x_1a_1, b_1y_1) and weight wt_1 is present as an HT in P^* , then it represents the exact same count of c HPs of its corresponding triple in P . This is exactly the number of HPs of the form (x_1a_1, b_1y_1) and weight wt_1 in G' . Its control bit β is set to 1.
- (c) $x_1 \in L(a_1, b_1y_1)$, $y_1 \in R(x_1a_1, b_1)$. Further, $(x_1a_1, b_1y_1) \in \text{Marked-Tuples}$ iff $c_{old} > 0$.
- (d) If $\beta(\gamma) = 0$ or $\beta(\gamma) = 1$ and there is an extension $x' \in L_j^*(x_1, y_1)$ that generates a centered LST in a level j , an LHT corresponding to $(x'x_1, b_1y_1)$ with weight $wt' = wt_1 + \mathbf{w}(x', x_1) \geq wt$ and counts equal to the sum of new paths represented by its constituents, is in H_f and P . A similar assertion holds for an extension $y' \in R_j^*(x_1, y_1)$.
- (e) The structure $RN(x_1, y_1, wt_1)$ contains a node b iff at least one path of the form $(x_1 \times, b_1y_1)$ and weight wt_1 is represented by a triple in $P(x_1, y_1)$. A similar assertion holds for a node a in $LN(x_1, y_1, wt_1)$.
- (f) The entry $DL(x_1, y_1)$ with weight wt_1 is updated to the current level.

Local Data Structures: for each level j , let c_j be the number of SPs of the form (x_1a_1, b_1y_1) and weight wt_1 centered in Γ_j and let $c_j(n)$ be the new ones discovered by the algorithm. Thus $c = \sum_j c_j$ and $c_{new} = \sum_j c_j(n)$. Then,

- (g) the value of $C_\gamma[j]$, where γ is the triple of the form (x_1a_1, b_1y_1) and weight wt_1 in $P(x_1, y_1)$, is c_j .
 - (h) If a triple γ of the form (x_1a_1, b_1y_1) and weight wt_1 is present as an HT in P^* , then $P_j^*(x_1, y_1)$ represents c_j paths. A link to γ in P is present in $dict_j$. Moreover, $x_1 \in L_j^*(a_1, y_1)$ (respectively $LC_j^*(a_1, y_1)$ if x_1 is centered in Γ_j). A similar statement holds for $y_1 \in R_j^*(x_1, b_1)$ (respectively $RC_j^*(x_1, b_1)$ if y_1 is centered in Γ_j).
3. For any key $[wt_2, x_2, y_2] \geq [wt, x, y]$, let G' contain $c > 0$ number of LHPs of weight wt_2 of the form (x_2a_2, b_2y_2) . Further, let c_{new} (resp. c_{old}) denote the number of such LHPs that are *new* (resp. not *new*). Here $c_{new} + c_{old} = c$. Then the tuple $(x_2a_2, b_2y_2) \in \text{Marked-Tuples}$, iff $c_{old} > 0$ and c_{new} paths have been added to H_f by some earlier iteration of the while loop.

Initialization and Maintenance for the invariant assertions above are similar to the proof of Lemma 6.

Termination: The condition to exit the loop is $H_f = \emptyset$. Because invariant assertion 1 maintains in H_f the first triple to be extracted and processed, then $H_f = \emptyset$ implies that there are no triples, formed by a valid left or right extension, that contain *new* SPs or LSPs, that need to be added or restored in the graph G . Moreover, because of invariant assertions 2a and 2b, every triple containing the node v , extracted and processed before $H_f = \emptyset$, has been added or restored with its correct count in the tuple-system. Finally, for invariant assertions 2c and 2h, the sets L, R and L^*, LC^*, R^*, RC^* for each level, are correctly maintained. This completes the proof of the loop invariant.

By Lemma 7, all the SP distances in G' are placed in H_f and processed by the algorithm. Hence, after Algorithm 7 is executed, every SP in G' is in its corresponding P^* by the invariant of Lemma 8. Since every LST of the form (xa, by) in G' is formed by a left extension of a set of STs of the form $(a \times, by)$ (Step 7 - Algorithm 7), or a right extension of a set of the form $(xa, \times b)$ (analogous steps for right extensions), and all the STs are correctly maintained and extended (by the invariant of Lemma 8), then all the LSTs are correctly maintained at the end of FF-FIXUP. This completes the proof of the Lemma. \square

References

1. D. A. Bader, S. Kintali, K. Madduri, and M. Mihail. Approximating betweenness centrality. In *Proc. of the 5th WAW*, pages 124–137, 2007.
2. H.-J. Bandelt and H. M. Mulder. Interval-regular graphs of diameter two. *Discrete Mathematics*, 50(0):117 – 134, 1984.
3. E. Bergamini and H. Meyerhenke. Fully-dynamic approximation of betweenness centrality. arXiv:1504.07091 [cs.DS], 2015.
4. E. Bergamini, H. Meyerhenke, and C. L. Staudt. Approximating betweenness centrality in large evolving networks. In *Proc. of ALENEX 2015*, chapter 11, pages 133–146. SIAM, 2015.
5. U. Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.
6. T. Coffman, S. Greenblatt, and S. Marcus. Graph-based technologies for intelligence analysis. *Commun. ACM*, 47(3):45–47, 2004.
7. A. Condon and R. M. Karp. Algorithms for graph partitioning on the planted partition model. *Random Struct. Algorithms*, 18(2):116–140, Mar. 2001.
8. C. Demetrescu and G. F. Italiano. A new approach to dynamic all pairs shortest paths. *J. ACM*, 51(6):968–992, 2004.
9. R. Geisberger, P. Sanders, and D. Schultes. Better approximation of betweenness centrality. In *Proc. of ALENEX 2008*, chapter 8, pages 90–100. SIAM, 2008.
10. O. Green, R. McColl, and D. A. Bader. A fast algorithm for streaming betweenness centrality. In *Proc. of 4th PASSAT*, pages 11–20, 2012.
11. D. R. Karger, D. Koller, and S. J. Phillips. Finding the hidden path: Time bounds for all-pairs shortest paths. *SIAM J. Comput.*, 22(6):1199–1217, 1993.

12. M. Kas, M. Wachs, K. M. Carley, and L. R. Carley. Incremental algorithm for updating betweenness centrality in dynamically growing networks. In *Proc. of ASONAM*, 2013.
13. N. Kourtellis, T. Alahakoon, R. Simha, A. Iamnitchi, and R. Tripathi. Identifying high betweenness centrality nodes in large social networks. *SNAM*, pages 1–16, 2012.
14. N. Kourtellis, G. D. F. Morales, and F. Bonchi. Scalable online betweenness centrality in evolving graphs. *IEEE Trans. Knowl. Data Eng.*, 27(9):2494–2506, 2015.
15. V. Krebs. Mapping networks of terrorist cells. *CONNECTIONS*, 24(3):43–52, 2002.
16. M.-J. Lee, J. Lee, J. Y. Park, R. H. Choi, and C.-W. Chung. Qube: a quick algorithm for updating betweenness centrality. In *Proc. 21st WWW Conference*, pages 351–360, 2012.
17. H. M. Mulder. Interval-regular graphs. *Discrete Mathematics*, 41(3):253 – 269, 1982.
18. M. Nasre, M. Pontecorvi, and V. Ramachandran. Betweenness centrality incremental and faster. In *MFCS 2014*, volume 8635 of *LNCS*, pages 577–588. Springer, 2014.
19. M. Nasre, M. Pontecorvi, and V. Ramachandran. Decremental all-pairs all shortest paths and betweenness centrality. In *ISAAC 2014*, volume 8889 of *LNCS*, pages 766–778. Springer, 2014.
20. J. W. Pinney, G. A. McConkey, and D. R. Westhead. Decomposition of biological networks using betweenness centrality. In *Proc. of 9th RECOMB*, 2005.
21. M. Pontecorvi and V. Ramachandran. Fully dynamic all pairs all shortest paths. <http://arxiv.org/abs/1412.3852v2>, 2014.
22. M. Pontecorvi and V. Ramachandran. Fully dynamic betweenness centrality. In *Algorithms and Computation - ISAAC 2015*, volume 9472 of *LNCS*. Springer, 2015.
23. J. S. R. M. Ramos and M. T. Ramos. A generalization of geodetic graphs: K-geodetic graphs. *Investigacin Operativa*, 1:85–101, 1998.
24. M. Riondato and E. M. Kornaropoulos. Fast approximation of betweenness centrality through sampling. In *Proc. of the 7th ACM WSDM*, pages 413–422. ACM, 2014.
25. S. E. Schaeffer. Survey: Graph clustering. *Comput. Sci. Rev.*, 1(1):27–64, Aug. 2007.
26. R. R. Singh, K. Goel, S. Iyengar, and Sukrit. A faster algorithm to update betweenness centrality after node alteration. In *Proc. of 10th WAW*, 2013.
27. N. Srinivasan, J. Opatrny, and V. Alagar. Bigeodetic graphs. *Graphs and Combinatorics*, 4(1):379–392, 1988.
28. M. Thorup. Fully-dynamic all-pairs shortest paths: Faster and allowing negative cycles. In *SWAT*, pages 384–396, 2004.